ENCE LIBRARY 18

DEC 1 8 1926

Vol. IX, No. 6 *PSYCHOLOGICAL REVIEW PUBLICATIONS* DECEMBER, 1926

Journal of Experimental Psychology

EDITED BY

MADISON BENTLEY, UNIVERSITY OF ILLINOIS

HOWARD C. WARREN, PRINCETON UNIVERSITY (*Review*)
JOHN B. WATSON, NEW YORK CITY (*Review*)
SHEPHERD I. FRANZ, UNIV. OF CALIF., SO. BR. (*Monographs*)
S. W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA (*Bulletin*)
WALTER S. HUNTER, CLARK UNIVERSITY (*Index*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY, Managing Editor

CONTENTS

Two Factors in the Work Decrement: EDWARD S. ROBINSON AND ARTHUR G. BILLS 415

Reading Reactions for Mathematical Formulæ: MILES A. TINKER..... 444

The Influence of Context upon Learning and Recall: SHUH PAN..... 468

Variability of Judgments on Musical Intervals: HELEN MORAN AND CARROLL C. PRATT 492

PUBLISHED BI-MONTHLY
FOR THE AMERICAN PSYCHOLOGICAL ASSOCIATION
BY THE PSYCHOLOGICAL REVIEW COMPANY
PRINCE AND LEMON STS., LANCASTER, PA.
AND PRINCETON, N. J.

Entered as second-class matter, July 1, 1920, at the post-office at Lancaster, Pa.

Psychological Review Publications

of the American Psychological Association

EDITED BY

HOWARD C. WARREN, PRINCETON UNIVERSITY (*Review*)
JOHN B. WATSON, 244 MADISON AVE., NEW YORK (*Review*)
SHEPHERD I. FRANZ, UNIV. OF CALIF., SO. BR. (*Monographs*)
MADISON BENTLEY, UNIVERSITY OF ILLINOIS (*J. of Exp. Psych.*)
S. W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA (*Bulletin*)
WALTER S. HUNTER, CLARK UNIVERSITY (*Index*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY, Managing Editor

WITH THE CO-OPERATION OF
MANY DISTINGUISHED PSYCHOLOGISTS

PSYCHOLOGICAL REVIEW

containing original contributions only, appears bimonthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 480 pages.

PSYCHOLOGICAL BULLETIN

containing critical reviews of books and articles, psychological news and notes, university notices, and announcements, appears monthly, the annual volume comprising about 720 pages. Special issues of the BULLETIN consist of general reviews of recent work in some department of psychology.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

containing original contributions of an experimental character, appears bimonthly, February, April, June, August, October, and December, the six numbers comprising a volume of about 480 pages.

PSYCHOLOGICAL INDEX

is a compendious bibliography of books, monographs, and articles upon psychological and cognate topics that have appeared during the year. The INDEX is issued annually in May, and may be subscribed for in connection with the periodicals above, or purchased separately.

PSYCHOLOGICAL MONOGRAPHS

consist of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The MONOGRAPHS appear at irregular intervals and are gathered into volumes of about 500 pages.

ANNUAL SUBSCRIPTION RATES

Review: $5.00 (Foreign, $5.25). Review and Bulletin: $9.50 (Foreign, $10.00)
Journal: $5.00 (Foreign, $5.25). Review and Journal: $9.00 (Foreign, $9.50)
Bulletin: $5.50 (Foreign, $5.75). Journal and Bulletin: $9.50 (Foreign, $10.00)
Any one of above with Index: $1.50 additional.
Review, Bulletin, and Journal: $14.00 (Foreign, $14.75).
Review, Bulletin, Journal, and Index: $15.00 (Foreign, $15.75).
Current Numbers: Review or Journal, $1.00; Bulletin, 60c; Index, $2.00
Psychological Monographs: $6.00 per volume (Foreign, $6.30).
Current Issues: prices vary according to size.

Subscriptions, orders, and business communications may be sent direct to the

PRINCETON, N. J. OFFICE

PSYCHOLOGICAL REVIEW COMPANY

FOREIGN AGENTS: G. E. STECHERT & CO., London (2 Star Yard, Cary St., W.C.), Paris (16, rue de Condé) ExP

Journal of
Experimental Psychology

VOL. IX., No. 6 DECEMBER, 1926

TWO FACTORS IN THE WORK DECREMENT

BY EDWARD S. ROBINSON

University of Chicago

AND

ARTHUR G. BILLS

University of Minnesota

I. THE GENERAL PROBLEM

Although the work decrement is one of the major phenomena of psychology, the experimental problems that it presents have not, in our opinion, received anything like the detailed attention that they deserve. If there is one truest generalization about the work decrement, it is a statement of its variability. Such a statement, of course, suggests the complexity of the principles underlying the losses in efficiency from relatively continuous work and, at the same time, points out an inevitable theoretical and experimental program. There is need for forgetting the ancient explanation of all sorts of decrements in terms of the depletion of a single supply of neural energy. And there is a further need for formulating the fundamental factors in the decrement, whether their nature be physical, biological, or psychological. Among others, Thorndike,[1] Dodge,[2] Myers,[3] and, more recently, one of the present writers,[4] have sought to establish theoretical

[1] E. L. Thorndike, *Educ. Psychol.*, 1912, 3, 1–140 (espec. 111–125).

[2] R. Dodge, The laws of relative fatigue, *Psychol. Rev.*, 1917, 24, 89–113.

[3] C. S. Myers, Conceptions of fatigue and adaptation, *ibid.*, 1925, 32, 1–16.

[4] E. S. Robinson, Principles of the work decrement, *ibid.*, 1926, 33, 123–134.

frameworks adequate to encompass the multifarious facts involved in the psychological work decrement. But experiment, we believe it fair to say, has lagged behind theoretical analysis. While theory has tended to see much more in the ordinary work decrement than refractoriness of tissue, experiment with typical psychological processes has been resorted to but little. Relations have been made out between the magnitude of the decrement and the type of efficiency considered, accuracy, speed, or satisfyingness,[5, 6] and between the magnitude of the decrement and the age and race of the workers.[7, 8] But these are only beginnings. There is some question, too, how near these studies have come to the heart of the problem. It was our purpose in the experiments reported in this paper to study some of the fundamental conditions of variability in the magnitude of the work decrement as it occurs in processes of a predominantly psychological kind.

II. The Factor of Homogeneity-Heterogeneity

We selected as our first factor for investigation the homogeneity of the work. In common thought, relatively homogeneous work is often considered to be more susceptible to 'fatigue' than work containing a greater variety of part-activities. Such work is, of course, most often characterized in terms of its subjective or affective nature, that is, in terms of its monotony, rather than in terms of its objective homogeneity. But in any event the factor of homogeneity, or monotony, seemed worth bringing under experimental control. Not only has it been assigned a position of importance by popular thought, but scientific considerations also suggest reasons why its importance should be expected. As Verworn [9] and others have shown, refractoriness of living tissue to stimu-

[5] E. L. Thorndike, Fatigue in a complex function, *Psychol. Rev.*, 1914, 21, 402–407.

[6] *Same*, The curve of work and the curve of satisfyingness, *J. Appl. Psychol.*, 1917, 1, 265–267.

[7] T. R. Garth, Mental fatigue during continuous exercise of a single function, *Arch. of Psychol.*, 1918, no. 26.

[8] *Same*, Racial differences in mental fatigue, *J. Appl. Psychol.*, 1920, 4, 235–244.

[9] M. Verworn, *Irritability, a physiological analysis of the general effect of stimuli in living substance*, 1913, 154–188.

lation repeated within certain time-limits is one of the most general of biological principles. That this same refractoriness is present throughout the range of psychologically described phenomena is too familiar a fact to need discussion. Now it is the broad principle of refractoriness of response that gives a basis for interpreting the influence of homogeneity of work upon the magnitude of the work decrement. The fewer the part-activities involved in continuous work the shorter will be the time-intervals between the successive stimulations of each of these activities and, within a given time-range, the more frequent will be those successive stimulations. As one of us has previously pointed out,[10] the degree of refractoriness is a function of both the recency and the frequency with which an activity is called upon to function. From this we should expect the continuance of relatively homogeneous work to be somewhat hindered by the factor of refractoriness.

It was our program, then, to select for experimental purposes a number of tasks differing from one another mainly in homogeneity and to compare the magnitudes of the decrements associated with continued work at these several tasks.

A. Experiment 1: Writing (Unbroken Sequences)

Procedure: The general activity utilized in our first experiment consisted in the writing of alphabetical sequences of letters. Three degrees of heterogeneity were secured by requiring the subjects to write such a two-letter sequence as *ababab*, etc., or such a three-letter sequence as *abcabc*, etc., or the six-letter sequence *abcdefabcdef*, etc.

In order that we might treat our three degrees of heterogeneity as quantitatively definable, we provided that, in the course of each cycle of experimentation, the two-letter work should include all of the letters involved in the three-letter and six-letter work, and likewise that the three-letter work should include all of the letters involved in the six-letter work. There were thus 3 varieties of two-letter work as follows:

ababab .....
cdcdcd .....
efefef .....

There were 2 varieties of three-letter work:

abcabc .....
defdef .....

[10] *Psychol. Rev.*, 1926, 33, 127–129.

But there was only the one variety of six-letter work:

abcdef . . . . . .

A single cycle of the experiment included 6 days. On each of these days a subject wrote one of the above sequences as fast as possible for a period of twenty minutes. In the course of the whole cycle he performed once under each of the six experimental conditions.

The subjects of the experiment—24 graduate students in psychology—were divided into groups of 6. Each subject in each group went through the conditions in his own particular order. Subj. 1 began with the first of the conditions as listed above, went through the second on the list on the second day, and so on. Subj. 2 began with the second of the listed conditions, then took the third, and finally, on the sixth day, went through the first. This scheme was carried through so that in the course of a cycle of the experiment each condition would appear equally often in each serial position. While this is not a perfect method of equalizing practice effects, it was quite adequate for our purposes. Since we ran the experiment for three successive cycles we had consistency of results as an additional check.

The writing was performed on sheets of yellow copy-paper, 13 x 8½ inches. The pencils used were fairly soft. As we have said, the subject was directed to write as rapidly as possible without making errors. Although there were occasional errors, these were not frequent enough nor regular enough to be of any importance. During the total writing-period of 20 minutes the experimenter gave an auditory signal—a light tap—at the end of each minute, whereupon the subject drew a vertical line through his writing at the point that he had then reached. This gave us a record of the amount of work done in each of 20 successive one-minute periods. In this first experiment, as in those to follow, all experimenting was done individually.

Results: Work curves were constructed for the three degrees of heterogeneity. Results from the 3 two-letter conditions were combined into a general two-letter work-curve. Likewise, a general three-letter work-curve was secured by combining the results from the 2 three-letter conditions. No combination was necessary in the case of the six-letter work, since it was represented by only one set of letters.

The results obtained from all 24 of our subjects were pooled. We have, however, kept our data separate by cycles. Since all of the subjects went through all of the experimental conditions 3 times we thus obtained data for 3 sets of work-curves. We believe that the consistency of these three sets of curves is, perhaps, the most important criterion of the reliability of this experiment.

Figures 1, 2, and 3 contain the curves of work of the three stated degrees of heterogeneity. Throughout the 3 cycles of the experiment the levels of the curves maintain the same general relationships. The two-letter work is invariably

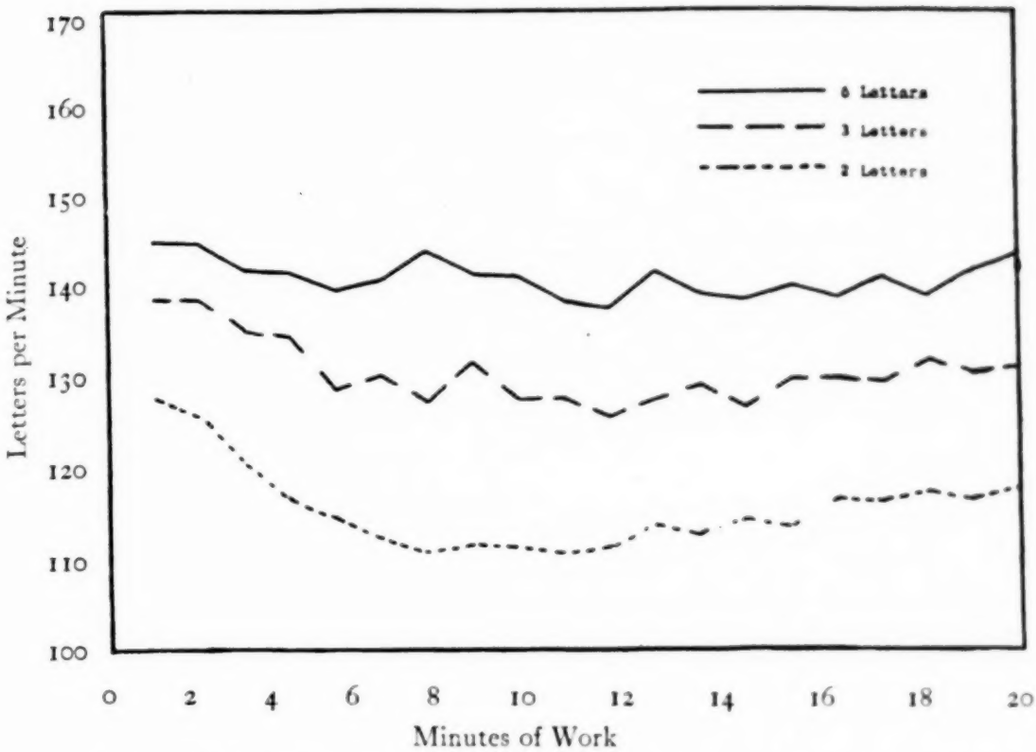

FIG. 1. Work-Curves for Cycle I of Experiment 1. Horizontal axis represents successive minutes of work. Vertical axis represents letters written per minute. Curves are averages for 24 subjects.

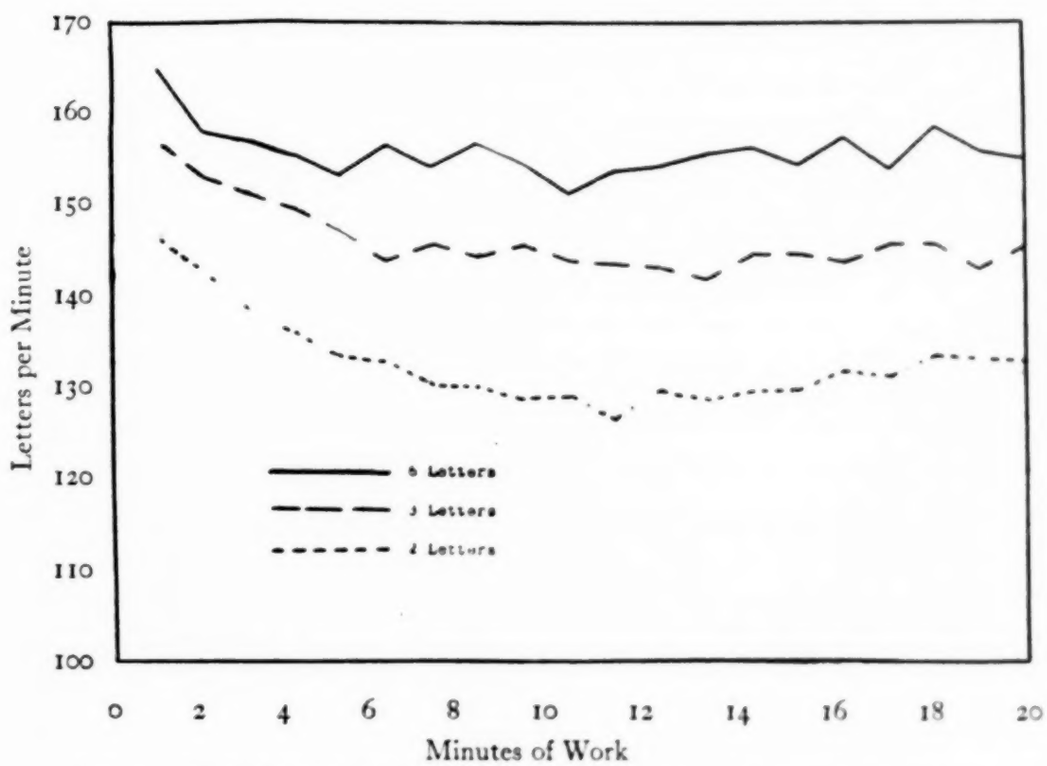

FIG. 2. Work-Curves for Cycle II of Experiment 1. See legend for Figure 1.

slower throughout than the three-letter work, and the three-letter work is invariably slower than the six-letter.

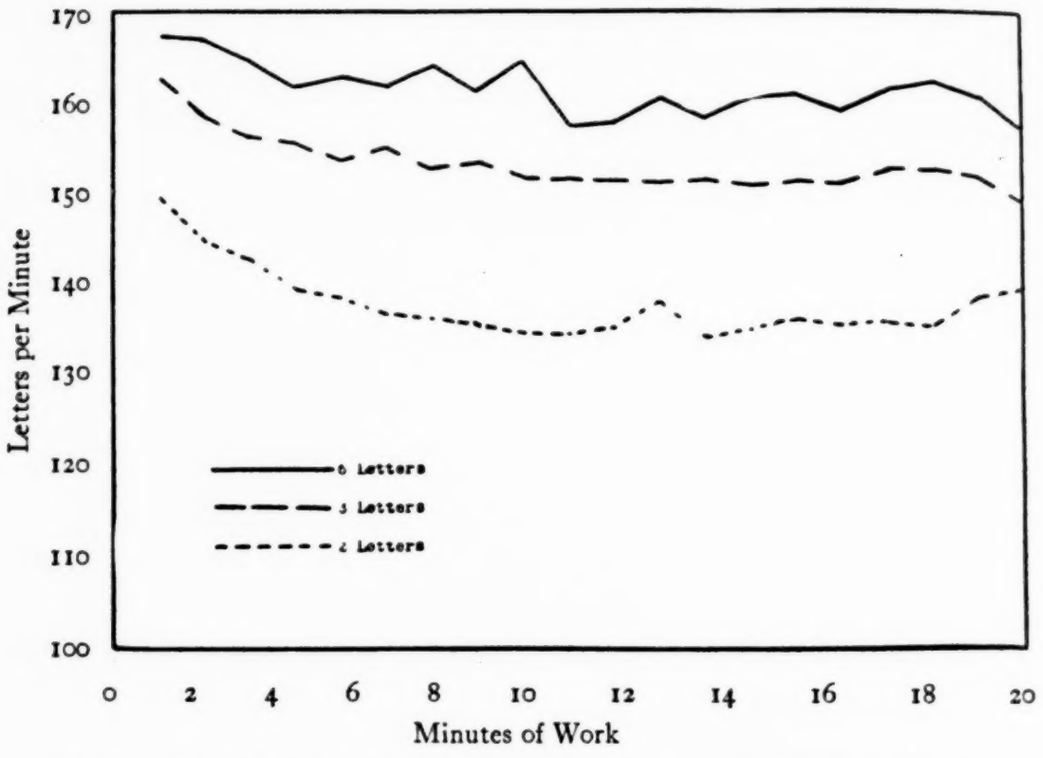

FIG. 3. Work-Curves for Cycle III of Experiment 1. See legend for Figure 1.

From casual inspection it seems that the curves for two-letter work are, in general, smoothest and those for six-letter work are, in general, most irregular. This is, in all probability, merely a statistical matter. Since there were 3 two-letter conditions and since there was only one six-letter condition, the two-letter curves are based upon 3 times as many observations as the six-letter curves. Being represented by two conditions, the three-letter work curves are based upon 2/3 as many observations as those for the two-letter work and upon twice as many as those for the six-letter work. And in smoothness these three-letter curves stand about as they should. Although that obtained from the first cycle is about as irregular as the six-letter curve, those obtained from the second and third cycles are about as smooth as the two-letter curves. There is no reason, then, to ascribe such differences in regularity as appear in our curves to anything except to the particular set-up of this experiment.

Our main interest is, naturally, in the decrements shown by these curves of work. A rough inspection reveals the fact that, throughout the experiment, the greatest decrement occurs in the most homogeneous work, i.e., that involving two letters. The difference between the three- and the six-letter work in this respect is not so clear.

In order to make more definite the facts of decrement, our

results were treated as follows. The differences in efficiency between the first and the tenth minute of work were determined for the three degrees of heterogeneity. These differences can be considered as measures of decrement magnitude. After inspecting our work-curves, we selected the tenth minute as representative of the maximal decrement-effect. Beyond the neighborhood of that point the curves either continue at about the same level or show an actual rise. Almost the same results were obtained by determining the decrement in terms of the difference between the first and the poorest minute of work.

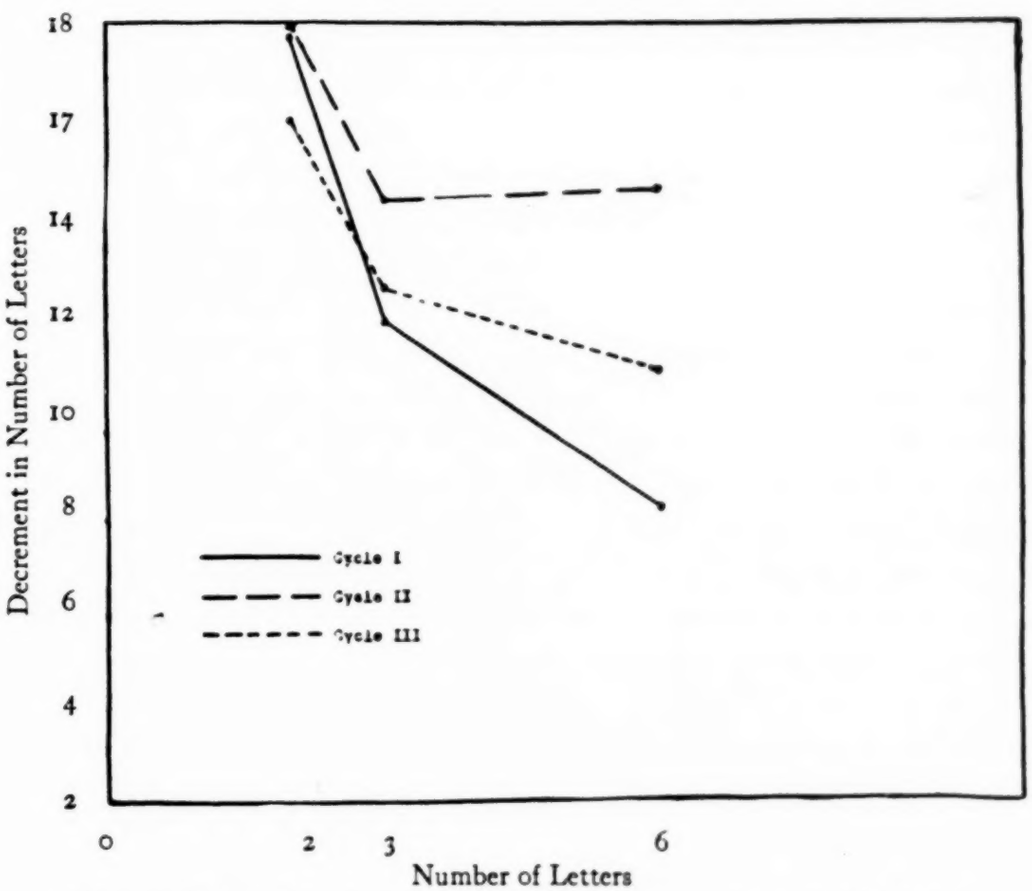

FIG. 4. Relationships between Degrees of Heterogeneity and the Work Decrements of Experiment 1. Horizontal axis represents degrees of heterogeneity in terms of number of letters involved. Vertical axis represents decrements in terms of difference between number of letters written in first minute and number of letters written in tenth minute.

In Fig. 4 there are plotted for the 3 cycles of the experiment the relationship between the heterogeneity of the work, as

measured in terms of number of letters involved, and the magnitude of the decrement, as measured in terms of the difference between the first minute's production and that of the tenth minute. All cycles agree in showing decelerated functions representative of the relationship between the heterogeneity of the work and its resistance to the decrement effect.

Our first experiment thus confirms our initial expectation that work containing few part-activities would show a more marked decrement than work containing more part-activities. The marked and consistent deceleration in this relationship, although we were not definitely expecting such a result, is hardly surprising. One would not anticipate a continuously linear relationship between resistance to decrement and heterogeneity of work. Even less would one anticipate a positively accelerated relationship between these two variables. The fact that we worked within a range where the more or less inevitable negative flexion happened to be manifest must, however, be looked upon as something of a happy accident.

The curves of Figs. 1, 2, and 3 represent the effects of continuous work upon *speed* of performance. Although no significant results were obtained in terms of relative frequency of right and wrong letters written, we noted early in the experiment that the continuation of work of this type did seem to affect the size of the writing. We determined for the entire experiment the average number of letters per line written during the successive minutes of work. The pooled results for all cycles and all subjects appear in Fig. 5.

Under all degrees of heterogeneity there is a marked increase in size of writing—decrease in number of letters per line—throughout the twenty-minute work period. But the magnitude of this effect does not bear any noticeable relationship to homogeneity-heterogeneity of work. We may dismiss it, therefore, as an interesting, but merely incidental finding.

Subjects' Reports: Although our subjects were not instructed to take an introspective, or process attitude, they were requested at the end of each sitting to report such incidental observations as they might have made. Several facts tand out consistently in these reports.

(1) The painfulness of the activity and the impulse to relax reached their maxima long before the end of the twenty-minute work period. After that the arm was somewhat numb and the work seemed to continue with a degree of automaticity.

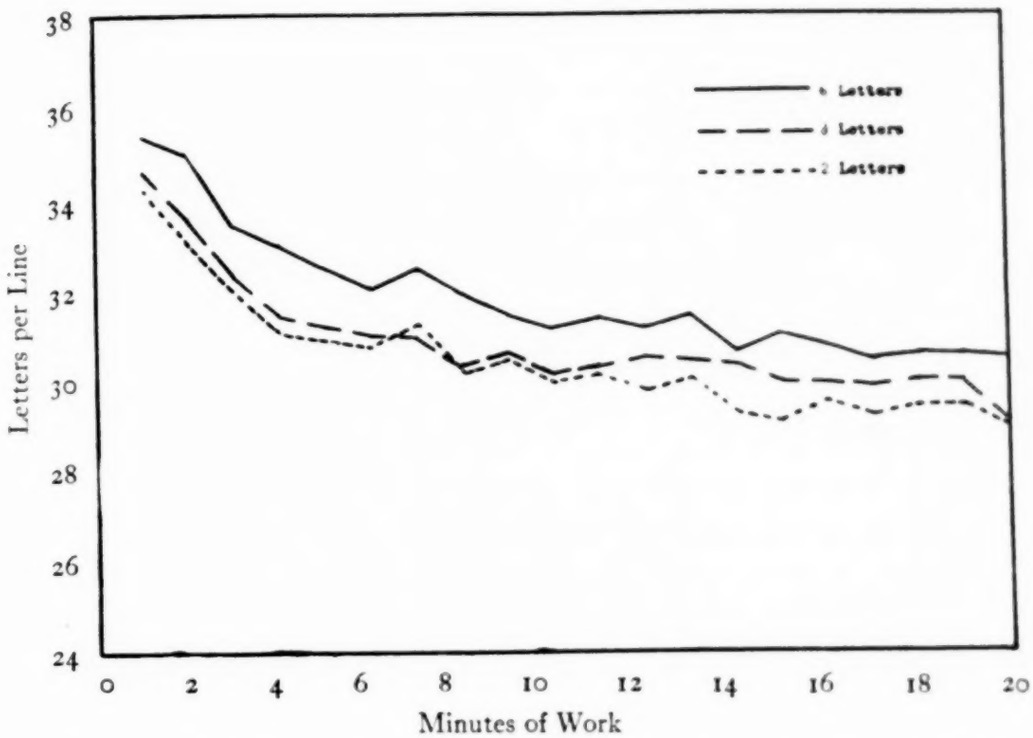

FIG. 5. Work-Curves for Experiment 1 showing Increasing Size of Writing during Period of Work. Horizontal axis represents successive minutes of work. Vertical axis represents number of letters per line. Curves are averages for all cycles and all subjects.

(2) When a 'blocking' occurred and it seemed as though the hand could not make the next letter, the pronunciation of that letter usually set the hand off to its proper task.

(3) In general the work went on best if the subject thought about something else. 'Blockings' frequently occurred when the subject became clearly aware of the writing.

B. Experiment 2: Writing (Broken Sequences)

Procedure: There was one element in our first experiment which slightly bothered us. Although we had found consistent differences in decrement between work of different degrees of heterogeneity, we had not isolated that relationship with

the completeness which might be desired. The tasks studied were carried out from the beginning at several levels of efficiency and there seemed to us a slight, if unlikely, possibility that the nature of our results might have been different had we been able to secure activities varying in homogeneity and decrement but approximately alike in initial efficiency. Our second experiment was an attempt to demonstrate the homogeneity-decrement in work performed with the same initial efficiency.

As a first step we asked ourselves why, independently of the development of any marked decrement, the subjects found it so much easier to write a six-letter sequence like *abcdefabcdef*, etc., than a three-letter sequence like *abcabc*, etc., or a two-letter sequence like *abab*, etc. Of course, there is the possibility that a decrement related to homogeneity had developed before the end of the first minute of work. But there is another possibility that seemed to us more credible. The writing contained alphabetical sequences and such sequences may be looked upon as firmly established habits. When, however, the subject came to the end of his 2 letters in the two-letter work, of his 3 letters in the three-letter work, and of his 6 letters in the six-letter work, he was forced to violate an alphabetical sequence by writing such a sequence as *ba*, or *ca*, or *fa*. Now such alphabetic violations might tend to block the smooth progress of the work, in which event those tasks containing the most frequent violations might be expected to be carried out at the lowest rate of speed. In the two-letter work these violations occurred between every other pair of letters, in the three-letter work they occurred between every third pair, and in the six-letter work between every sixth pair. Thus the frequency of the alphabetic violations does, in fact, correspond inversely to the general level of efficiency. We therefore set about to alter our experiment along the lines indicated by these considerations.

The first procedure that occurred to us was that of utilizing only non-alphabetical sequences. This plan was rejected because of the fact that, under this condition, there would be a much larger learning factor involved in the writing of six-

letter sequences than in the writing of two-letter sequences. The plan adopted was that of having the subject break his writing—that is, leave such a space as is ordinarily left between words—at the end of each two letters. In the first experiment, the writing ran continuously from letter into letter with breaks occurring only at the end of the lines. We felt that a sequence like *abcdef*, if written *ab cd ef*, would have little advantage over a sequence like *ab ab ab*. The breaks between pairs of letters might be expected to eliminate the difference which would exist between an alphabetical and a non-alphabetical transition.

In this second experiment we studied work of only two degrees of homogeneity-heterogeneity, that is, two-letter work and six-letter work. Following the plan of the first experiment, we had the same letters represented in both degrees of heterogeneity. We had the following kinds of two-letter work:

ab ab ab, etc.
cd cd cd, etc.
ef ef ef, etc.

The six-letter work consisted in writing:

ab cd ef ab cd ef ab cd ef, etc.

The same general plan of equalizing the effects of practice was employed. The same 24 subjects served. Each one went through all of the above conditions twice. Because of the fact that in the first experiment practically all of the decrement obtained was obtained within the first 10 minutes, we limited the total working period to this time. Marks were made across the writing every half-minute instead of every minute. Experiments were made individually and a subject worked under only one condition on any day.

Results: Figs. 6 and 7 show the curves for the two-letter and six-letter work secured by combining the results of all subjects for each of the two cycles of the second experiment. As in Experiment 1, generalized two-letter curves were obtained by combining the results from the 3 two-letter conditions.

This second experiment strikingly supports the first in that the more heterogeneous work is much more resistant to the decrement effect. But the really important fact brought out

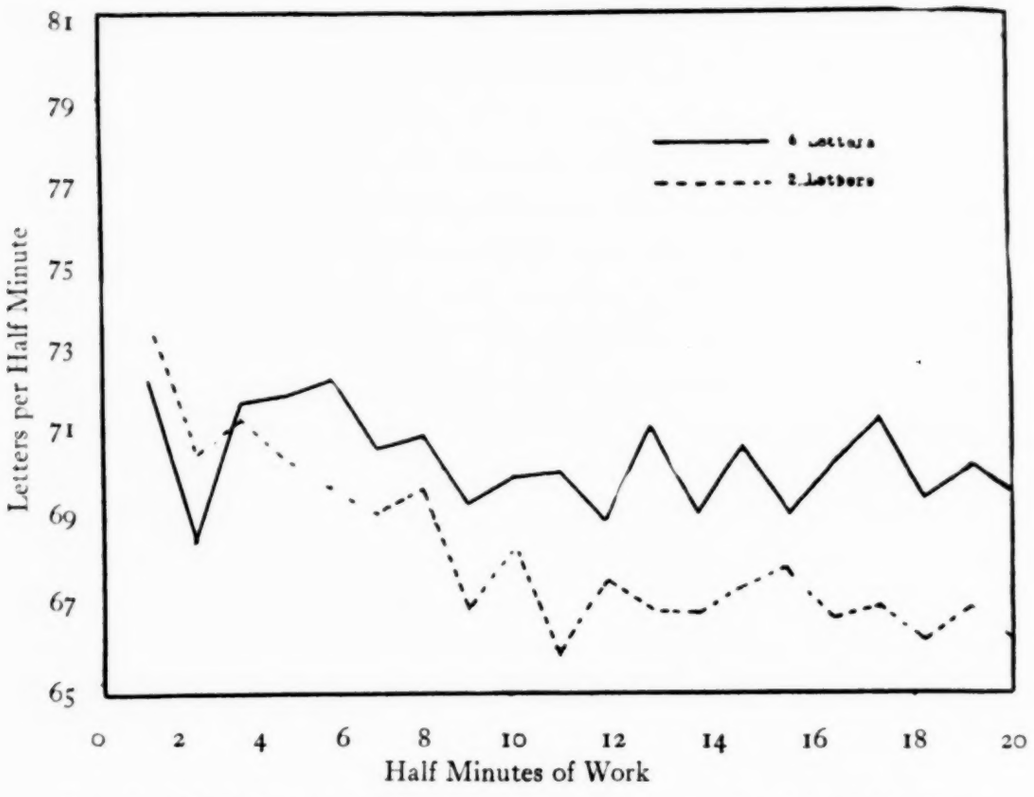

FIG. 6. Work-Curves for Cycle I of Experiment 2. See legend for Figure 1. Curves are averages for 24 subjects.

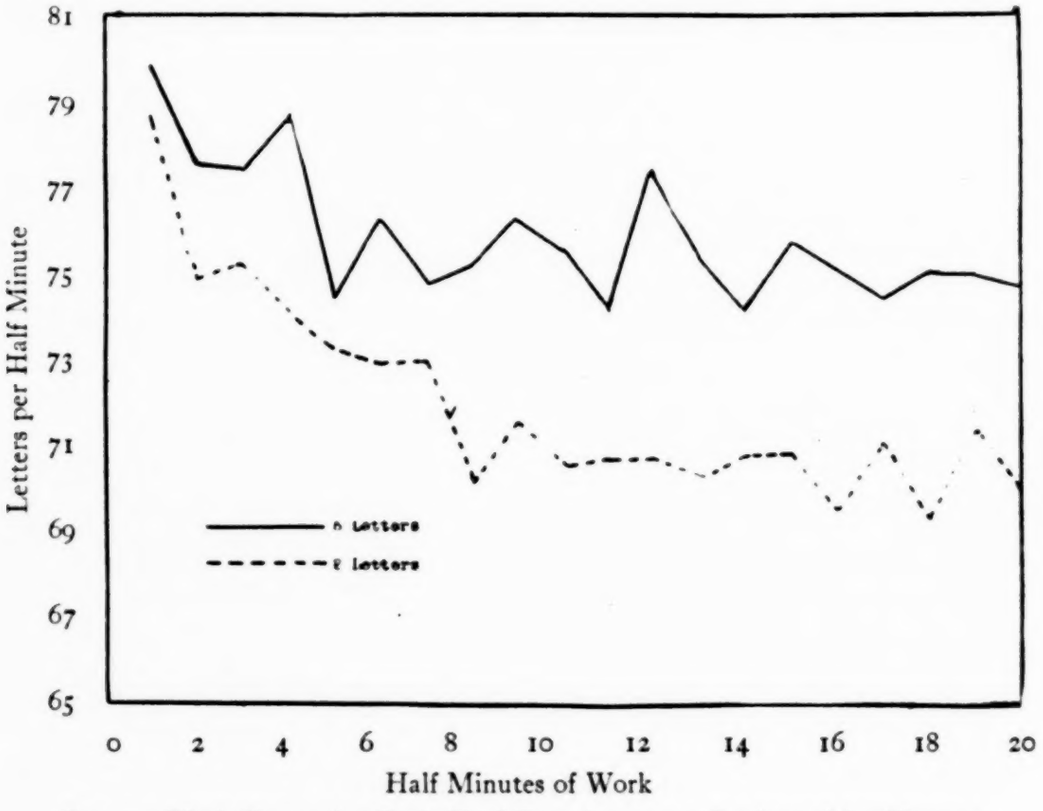

FIG. 7. Work-Curves for Cycle II of Experiment 2. See legend for Figure 1.

by this second experiment is that this difference in decrement can be obtained from activities of the same or of similar initial efficiency.

A word might be said as to whether the similarity in initial efficiency achieved in our second experiment was really achieved by reducing the effects of alphabetic violations in the tasks of unlike heterogeneity. By the time of the second experiment, the subjects had had considerable practice in work of this general type and it is conceivable that this practice might have somehow reduced the difference in initial efficiency between the different tasks. If this were the case, however, we should expect to find evidence for it in the successive cycles of the first experiment. But throughout that first experiment the differences in initial efficiency under the different degrees of heterogeneity show little or no evidence of decreasing with practice. Thus it seems probable that our introduction of the breaks after every two letters was the potent factor in obtaining similar initial efficiency from the tasks employed. But however that may be, we have demonstrated that decrement can be increased by increasing the homogeneity of work, even if the initial level of efficiency of that work be left unchanged.

C. Experiment 3: Letter-Naming

Procedure: It was the primary purpose of our third experiment to test the generality of the conclusions from the first two. We wanted to see whether the same type of correlation between homogeneity and magnitude of decrement might be expected to hold for work differing considerably from the letter-writing which we had thus far utilized. Of course, we were limited in the work that might be studied, because it was necessary to select a number of tasks which, while varying in homogeneity, should vary in some definable way. And aside from the matter of homogeneity the tasks needed to be highly similar. The type of activity finally hit upon was suggested to us by the familiar color-naming test, which utilizes a card upon which patches of several different colors are spread at random. The subject goes through the patches, naming

each color as he comes to it. For the purposes of our study we found it desirable to substitute letters for colors. Otherwise our procedure was practically that of the color-naming test. This gave us an activity which, as compared with the letter-writing, was more perceptual than overtly motor.

The reason why we used letter-naming rather than color-naming can be explained in terms of a secondary goal that we had set ourselves. In the first of the letter-writing experiments we had had three degrees of homogeneity-heterogeneity. This enabled us to establish within a very restricted range a markedly and consistently decelerated function representative of the relationship of the homogeneity factor to the magnitude of the work decrement. We hoped that in the third of our experiments we might do more than test the generality of our earlier findings. We hoped that, by studying—in the case of some additional activity—a greater number of grades of heterogeneity, we might secure a more comprehensive picture of the relationship of this type of variability to variability in decrement. By using 2, 3, 4, 5, or possibly 6 different colors in the color-naming test one could study the effects of corresponding degrees of heterogeneity, but the use of more than 6 colors would run into difficulties connected with the general familiarity of the subjects with the nomenclature of color. There are, however, 26 letters whose names may be assumed to be well established for all adult subjects. Thus it was that, in order to secure as wide a range of homogeneity-heterogeneity as possible, we substituted letters for colors.

At each sitting a subject was presented with a cardboard, 12 inches square. Upon this board there were 10 rows of $\frac{1}{4}$-inch gummed letters, 10 letters to a row. The subject was instructed to read through the card as rapidly as possible. Since every subject would have adopted some sort of rhythm for this work, we directed that the reading be in trochaic rhythm. Time was recorded for each reading of the card. No error score was recorded. As a matter of fact, there were only a few scattered cases of misnaming throughout the entire experiment. Inefficiency appeared almost always simply as a temporary inability to call a given letter by any name. The subject read through the same card 20 times without rest between readings. Only one such performance was engaged in by a subject on a single experimental day.

In all there were five kinds of cards, each kind representing a degree of homogeneity-heterogeneity. One kind contained 2 different letters, another 4, another 8, another 16, and another 24. These letters appeared, of course, in random order. We had as subjects 18 graduate students in experimental psychology who were unaware of the results of our first two

experiments and also unaware of the specific purpose of the experiment in which they were working. All of these 18 subjects went through all 5 conditions of the experiment 4 times. In order to prevent the domination of the effects of practice the subjects went through the 5 conditions in a variety of orders. Since the number of subjects was not an even multiple of the number of conditions, we were unable to have every experimental condition occur an equal number of times in every practice-order. That our less precise method of minimizing the effects of practice was adequate will, however, be apparent when we come to consider the consistency of our results throughout the several cycles of the experiment. In order to minimize accidental differences in the difficulty of various letters and arrangements of letters, we utilized in the experiment as a whole 16 different two-letter cards, 12 different four-letter cards, 8 different eight-letter cards, and 4 each of the sixteen- and twenty-four-letter cards. We were, of course, assuming that the more letters there were involved in the make-up of two cards of the same type, the less likely would there be important chance differences in difficulty between them. Our general set-up would have been nicer had we been able to represent in each cycle of the experiment and in the work of each subject the same letters in the two-letter work as were involved in the most heterogeneous work. But our ability to utilize such a scheme in our first experiments was based upon the fact that we were there studying a very limited range of homogeneity-heterogeneity. To have followed the same plan in this third experiment, which was designed to cover four times the range of homogeneity-heterogeneity covered by the first two, would have been quite impracticable.

Results: Figs. 8, 9, 10, and 11 contain the work-curves secured during the four cycles of this experiment. Each curve represents the pooled performances of the entire group of 18 subjects. From a superficial survey of these curves it is evident that the most general of our earlier findings are here confirmed. That is to say, the decrements are most marked where the elements in the work are relatively few and least marked where they are relatively numerous. There are,

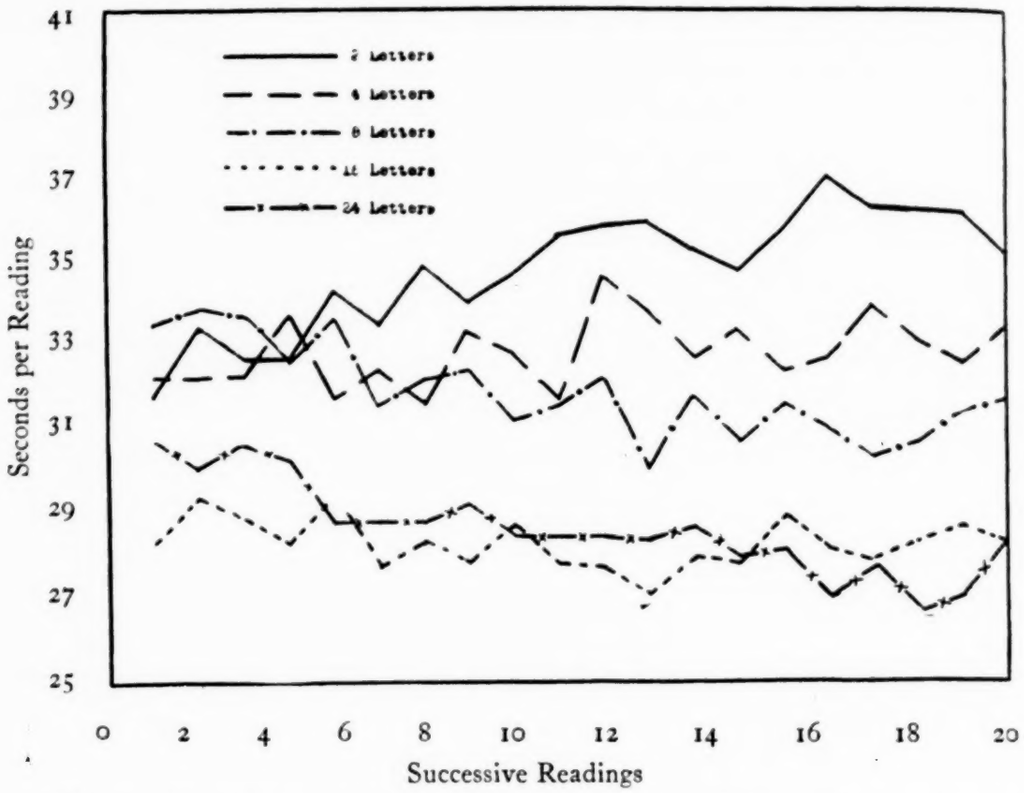

FIG. 8. Work-Curves for Cycle I of Experiment 3. Horizontal axis represents successive readings of card containing 100 letters. Vertical axis represents seconds per reading. Curves are averages for 18 subjects.

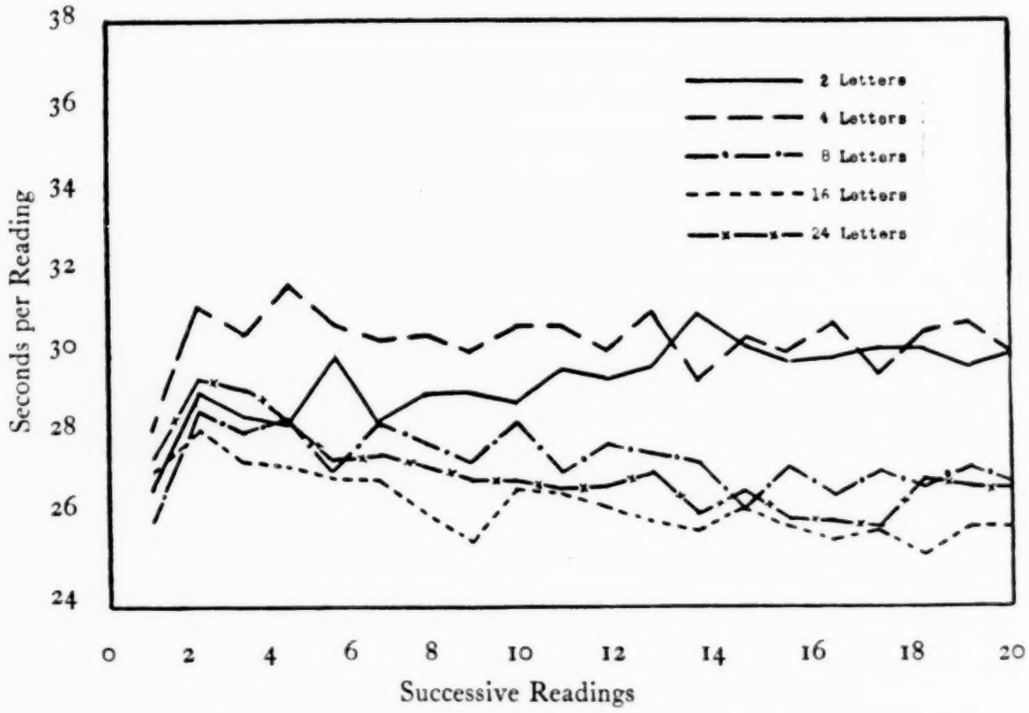

FIG. 9. Work-Curves for Cycle II of Experiment 3. See legend for Figure 8.

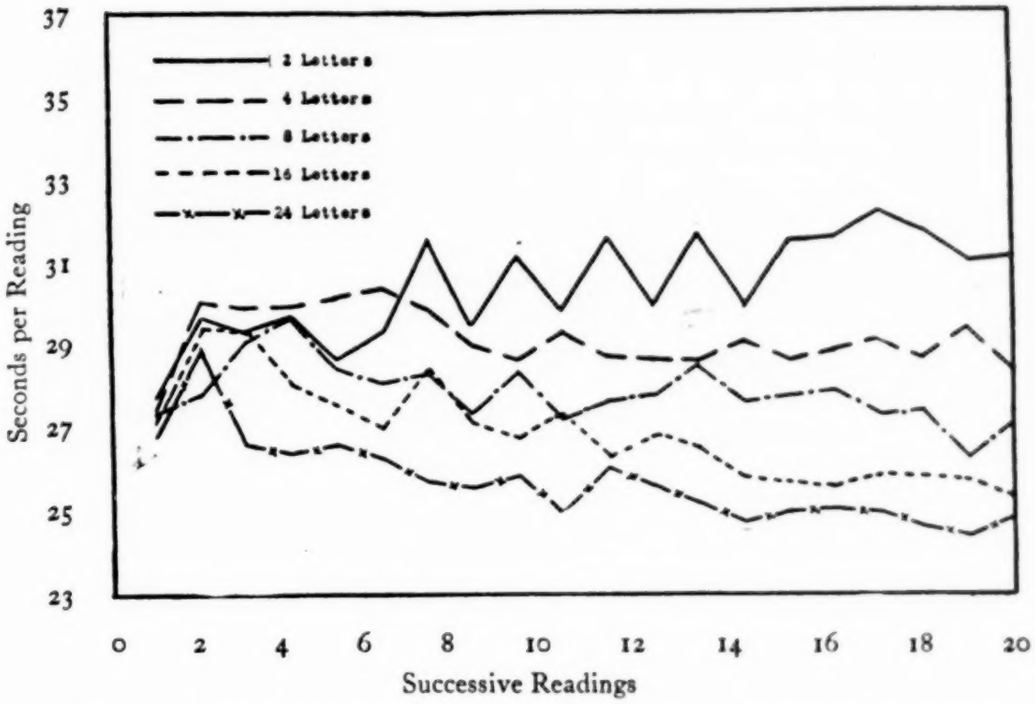

FIG. 10. Work-Curves for Cycle III of Experiment 3. See legend for Figure 8.

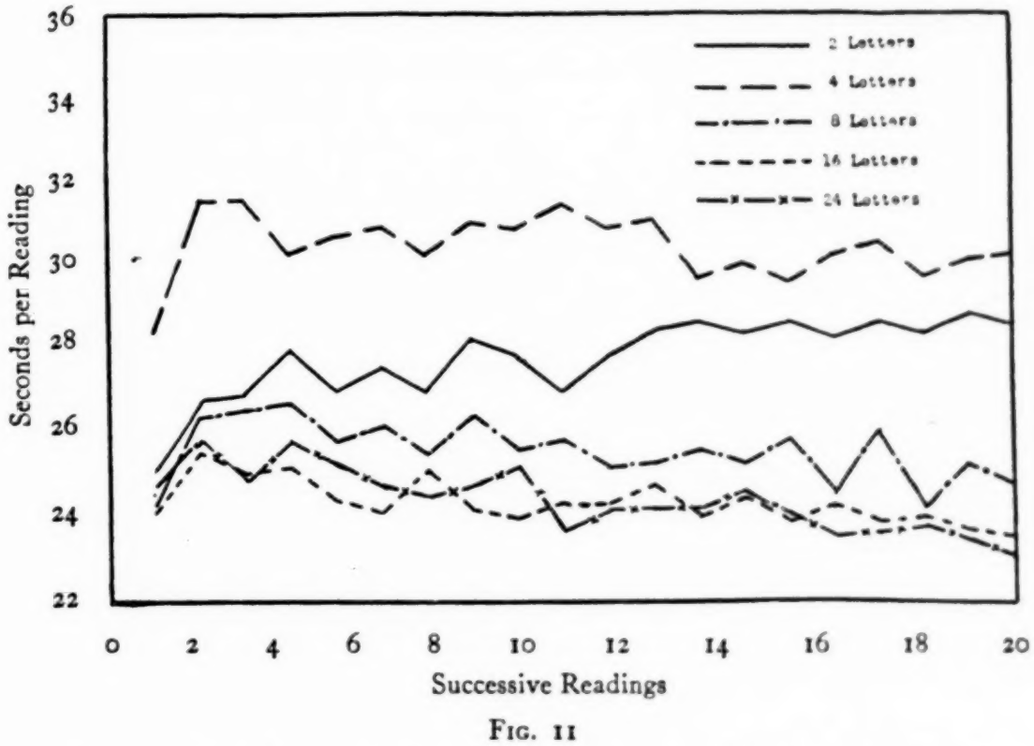

FIG. 11

FIG. 11. Work-Curves for Cycle IV of Experiment 3. See legend for Figure 8.

nevertheless, a good many irregularities which render more exact conclusions from inspection difficult. In the first cycle the two more heterogeneous types of work have a much greater initial efficiency than the three more homogeneous types. After the first cycle the various types of work have about the same initial efficiency, though there is an outstanding exception in the case of the four-letter work in the fourth cycle. We must there have drawn one or more very difficult combinations of letters.

Most of the curves contain a brief *initial spurt* (early high level of efficiency from which there is a relatively sudden falling off). There is reason to believe that the abrupt drops from the period of spurt are to be considered as fairly independent of the more general falling off which we are accustomed to find in most curves of work. The initial spurt is about as marked in the curves which, when considered as wholes, show a minus decrement as it is in those curves which show a decrement throughout their entire length. We are at a loss to explain the fact that an initial spurt of this definite character appears in our third experiment but does not appear in our first two. Chapman and Nolan [11] have pointed out that initial spurt may be present without being detected, if the divisions of the work-curve are too few. If the first point on the curve represents the average level of efficiency for a considerable time-period, an initial spurt may be contained within that period and yet be obscured by virtue of the fact that it consumes only a small fraction of the period represented by that single point on the work-curve. That does not seem to be the case here, however, because the recording of production in the first two experiments was not in terms of periods much larger than those involved in the third experiment. There is some basis for explaining why the second reading of the cards should so generally have been slower than the first. This matter simply looks like a type of refractoriness that one might expect as resulting from early repetition. But why efficiency so often improves after that second reading and why

[11] C. J. Chapman & W. J. Nolan, Initial spurt in a simple mental function, *Amer. J. Psychol.*, 1916, 27, 256–260.

there is no early abrupt decrement in the curves for letter-writing we are quite unable to say.

The fact that the curves of the more heterogeneous work show an actual improvement in efficiency after the brief period of initial spurt is undoubtedly related to a tendency on the part of the subjects to learn a card as they repeated it. Indeed, some of the subjects reported that there was a memorization factor present in their performance. That this increase in efficiency is not a general practice effect for this kind of work is indicated by the persistence of the increase throughout the later cycles of the experiment, where we may assume that general practice effects had become somewhat diminished.

But we have yet to touch on that phase of our results which is our main interest. We refer to the more exact relationship between homogeneity-heterogeneity and the magnitude of the work decrement. We determined the decrements in the

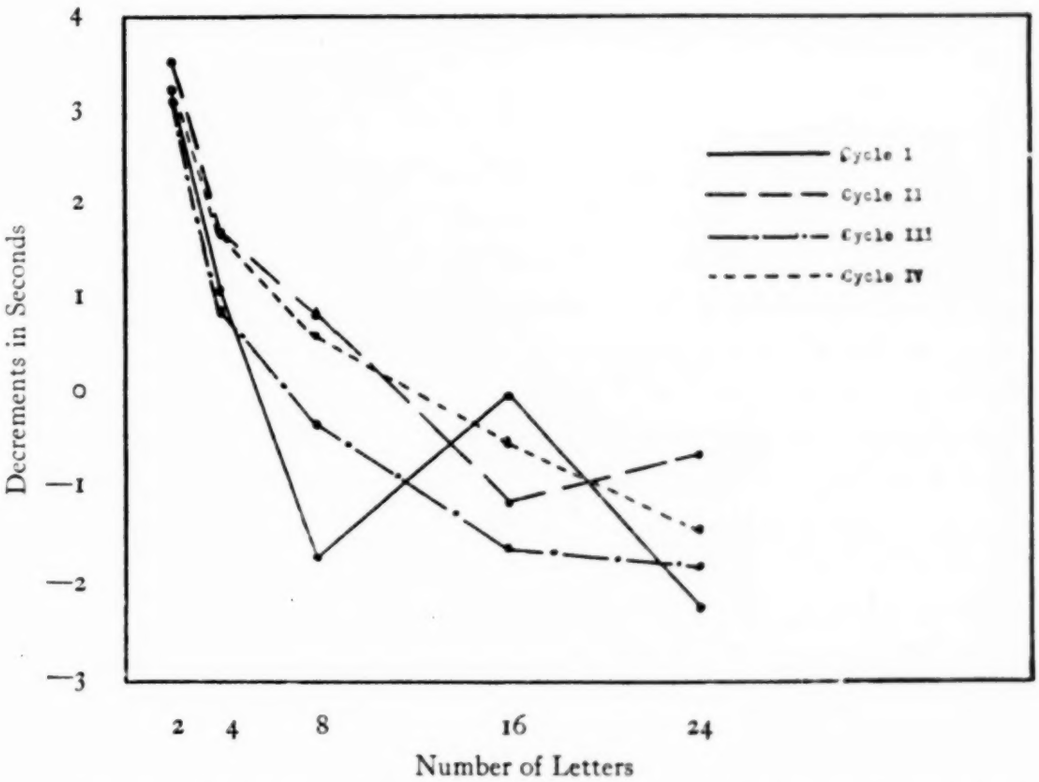

FIG. 12. Curves showing Relationships between the Degrees of Heterogeneity and the Work Decrement of Experiment 3. Horizontal axis represents degrees of heterogeneity in terms of number of letters involved. Vertical axis represents decrement in terms of difference in time required for first reading of card and for twentieth reading.

curves of the four preceding figures by getting the difference between the rate of the first reading of a card and its last read-

ing. We determined the decrement from these two end-points of the work-curves because, in those curves which showed a decrement most markedly, there was no indication of that decrement tending to come to an end before the end of the total recorded period of work. In Fig. 12 we have plotted for the four cycles of the experiment the relationships between the work decrement and the heterogeneity of the work.

The functions secured are quite unambiguous. There is generally a regular, decelerated relationship between heterogeneity and resistance to the decrement effect. In each of the first two cycles there occurs a first-order inversion of this relationship, but even in these cases the trend of the functions is clearly as we have stated. Thus, in an experiment representing a much wider range of homogeneity-heterogeneity than was present in the letter-writing, we are able to confirm our earlier findings.

At this point there is one statement about this homogeneity-heterogeneity matter that needs to be made. We do not for one moment assume that wherever different tasks show differences in homogeneity-heterogeneity there will be differences in decrement corresponding to the law discovered in our first three experiments. We are merely concluding that when decrement varies, homogeneity-heterogeneity variation *may* be the principal causal factor. The fact that the relationships found by us have such marked negative flexions suggests that, after a certain fairly low degree of heterogeneity is passed, further increases in this factor might have little influence upon the decrement. Dodge [12] introduced his concept of *relative fatigue* in order to explain decrements whose magnitudes seemed fairly independent of refractoriness brought about by sheer recency or frequency of functioning of the activity in question. As one of us has shown,[13] we can hardly question the importance of other principles than those of refractoriness brought about by recency and frequency of functioning. Our three experiments should be taken then merely as demonstrating that in work of a typically psychological sort homo-

[12] *Psychol. Rev.*, 1917, 24, 101–113.
[13] *Psychol. Rev.*, 1926, 33, 129 ff.

geneity-heterogeneity (that is to say, recency and frequency of functioning—see p. 416) is capable of being a notable determiner of degree of decrement and that the importance of this factor is probably much greater where the tasks concerned are, as tasks go, of a relatively homogeneous sort.

Subjects' Reports: In the third of our experiments the subjects were not directed to take a process attitude and it may be assumed that the attitude predominating during the work was objective. At the end of each day's performance we did, nevertheless, call for reports. There are certain observations which were made so frequently that they must be accorded some general significance.

(1) Eleven of the subjects noticed that the letter-naming went on much more smoothly while they were engaged in apparently irrelevant trains of thought. These ideational activities seem to have been of the order of fairly concrete fantasies. This fact raises a question regarding the nature of the ideational process. While it is interesting enough that the letter-naming went on most effectively while the subject was engaged in thinking of remote matters, it is more interesting that elaborate processes of ideation should go on simultaneously with a continuous visual-vocal process irrelevant to that ideation. Evidently such visual images as these subjects had were not mere visual meanings of ocular movements, and evidently subvocal movements had no vital rôle in these ideational processes.

(2) It was found in the writing experiments that "blockings" were most often relieved by vocalizations. A similar phenomenon was reported in the letter-naming experiment by the majority of the subjects. In this case, however, the blockings were relieved by a wide variety of body and limb movements bearing no obvious relationships to the letters the reading of which they facilitated.

III. The Factor of Competition

Our first three experiments have shown how homogeneity-heterogeneity operates, when it does operate, as a principal determiner of the magnitude of the work decrement. The

next logical step appeared to be that of seeking a better understanding of the *limitations* of the influence of this factor. As we have already said, the fact that the functions obtained in the first and third experiments show such marked decelerations suggests that, after a certain minimal degree of heterogeneity has been passed, further increases in heterogeneity may not be associated with changes in the magnitude of the decrement. It also occurred to us that, if there are other determiners of decrement, the presence of one of them might at times be so pronounced as to obscure the effects of variations in homogeneity-heterogeneity. We let this second possibility set our next experimental goal, in that we sought to arrange an experiment in which the results of our earlier work might find an exception;—an experiment in which the more homogeneous work might show a decrement no greater than that shown by less homogeneous work. Such an experiment, we felt, would add to our knowledge of the limits of applicability of our earlier conclusions and, also, it might illustrate the importance as a decrement determiner of some other principle than that pertaining to the homogeneity of the work.

But where were we to find tasks of variable homogeneity the decrements of which might be expected to be dominated by a factor that would bring about the larger decrements in the more homogeneous work? This general possibility came to mind. The susceptibility of an activity to decrement might easily be a function of the nicety with which the part-activities making it up are balanced. If a total activity is composed of a number of delicately integrated part-activities, any disturbing factor, such as too frequent repetition of operation, might bring about a relatively profound loss of efficiency in the activity as a whole. Now there must be cases where, among acts of the same general type, those containing the most numerous part-activities are, by virtue of that very fact, more delicately integrated than those containing fewer part-activities and, therefore, more susceptible to the decrement effect. Of course, we do not mean that activities containing many part-activities will always be more delicately integrated and, therefore, more susceptible to disturbance, but merely

that more complicated acts are *often* more delicately set up than similar acts of greater simplicity. It is clear that in our first three experiments there was little reason to assume more delicacy of integration for the complex activities than for the simpler ones. But it seemed reasonable to hope that we might find a form of activity in which greater complexity might mean greater delicacy of adjustment among part-activities and thus greater susceptibility to such a disturbing factor as excessively frequent repetition.

It is a recognized principle that an act that exists always in close competition with certain other acts may suffer from too frequent repetition largely because a relatively slight loss in the efficiency of the act itself may be enough to favor a competing act, which may in turn block off the original act entirely.[14, 15] Thus where the simpler principles of refractoriness would explain only a slight decrement, the presence of the factor of competition may make the actual decrement very marked. When, therefore, we think of the possibility of more heterogeneous work showing a more marked decrement (where greater heterogeneity means greater delicacy of integration) we may perhaps assume that the delicacy of integration is itself to be thought of largely in terms of the intercompetition of part-activities. In this case we may rephrase our aim as one of discovering a set of tasks in which, because complexity indicates more possibilities for competition, heterogeneity may be expected to be a positive rather than a negative condition of decrement.

D. Experiment 4: Typing

Procedure: The tasks finally selected for study involved simple finger movements. The subject tapped with the first, second or third finger, or successively with two or three of these fingers upon the keys of a typewriter. The tapping was done with the better hand. The only keys used were those in the central portion of the bottom row of the key board; that is, *v*, *b*, and *n*. The margin-stops were set to allow 50 letters per line. Twenty lines were written at a single performance.

[14] *Psychol. Rev.*, 1917, **24**, 102, 105.

[15] *Psychol. Rev.*, 1926, **33**, 129–131.

The time was recorded for each successive line. Errors were recorded from the typed copy.

We considered the various tasks as falling into three homogeneity-heterogeneity classes, according to the number of fingers involved. There was one-letter work; this consisted of writing *vvvvv*, etc., with the first finger, *bbbbbb* with the second finger, or *nnnnnn*, etc., with the third finger. There was two-letter work; this consisted of writing *vbvbvb*, etc., with first (index) and second fingers, *bnbnbn*, etc., with the second and third fingers, or *vnvnvn*, etc., with the first and third fingers. There was three-letter work; this consisted of writing *vbnvbn*, etc., with the first, second, and third fingers. Conditions were properly reversed for the few left-handed subjects.

In all of the writing tasks the subjects lifted the finger a slight distance from the key at the end of each stroke. Otherwise the one-finger writing might have consisted of single strokes differing radically in character from those used where the strokes were made successively by different fingers. Of course, we could hardly be sure that the stroking of the first finger, when only that finger was being used, was identical in character with the stroking of that finger when it was used alternately with one or two other fingers. This was not really necessary, however, for the purposes of our experiment. All that was necessary was that the one-finger work be more homogeneous —that it involve a greater amount of repetition of part-activities per letter written— than the two- and three-letter work, and that the two-letter work be more homogeneous than the three-letter work. We believe that these conditions were fulfilled.

The writing was done on a late model Royal typewriter which was chosen because on this machine a prominent lever makes possible the rapid return of the carriage. Since we were studying continuous work, we naturally wanted as short a pause as possible between lines.

The subjects were instructed to work as rapidly as they could without letting the keys interfere with each other. This interference factor was absent from the one-letter work, but that was a negligible difficulty. It would tend to allow greater speed in the simpler work and that would tend to produce a more pronounced decrement in such work. If, despite this fact, the more heterogeneous work showed an equal or greater decrement, our expectations would be fulfilled under unusually rigorous conditions.

The subjects were 21 students from a graduate course in experimental psychology. They differed widely in typing experience, but since all subjects were represented under all conditions of the experiment we saw no reason to differentiate between them.

As indicated, there were in all 7 writing tasks, 3 one-finger tasks, 3 two-finger tasks, and 1 three-finger task. The subjects were divided into 3 groups of 7 so that within each group the different subjects could all begin and end a cycle on different conditions. All subjects went through all conditions

6 times. On any single day a subject went through the same condition twice, with a rest of about 10 minutes between performances. Not knowing how much the second performances of a given day might be affected by the first, we kept the records of these first and second performances separate.

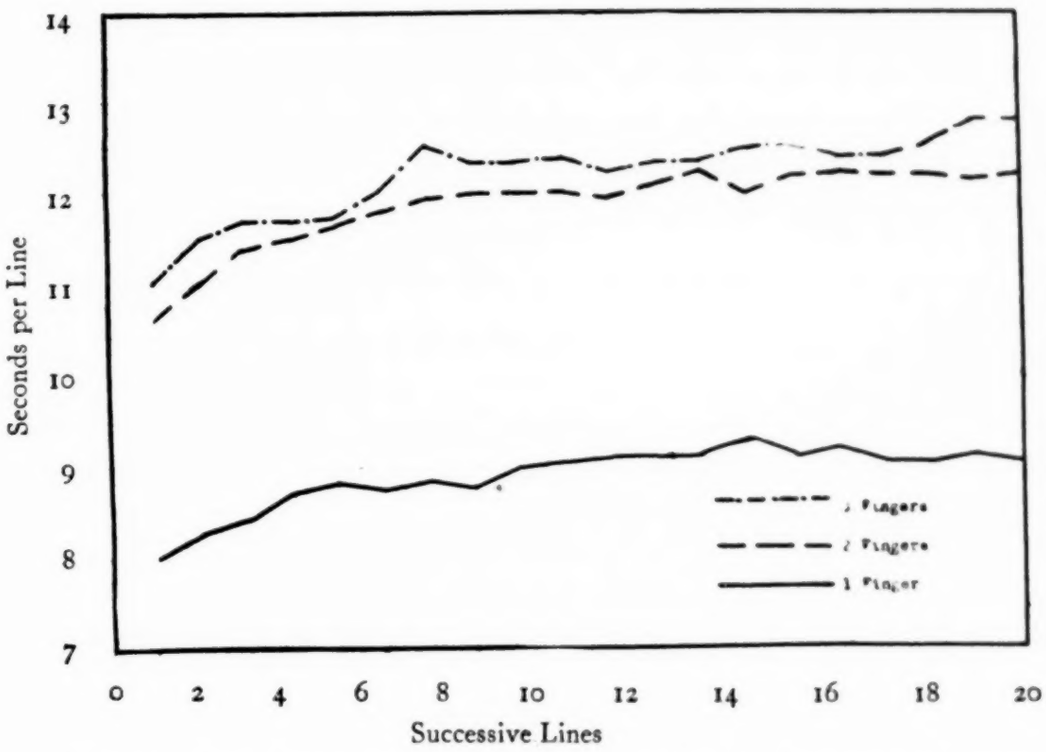

FIG. 13. Work-Curves for Experiment 4. First Trials. Horizontal axis represents successive lines written. Vertical axis represents speed in terms of seconds per line. Curves are averages for 21 subjects and 3 cycles.

Results: In Fig. 13 we have the work-curves for the three types of tasks, where those tasks were the first performed on an experiment-day. We have pooled the results of all subjects and all cycles of the experiment. The greater irregularity of the three-finger curve is almost surely a statistical matter. It represents the pooling of only one-third as many observations as were involved in making up each of the other two curves, because there was in the experiment only one form of three-letter work.

It is plain from inspection that the decrement is at least as great for the more heterogeneous tasks as for the most homogeneous. From the first to the twentieth line there is a loss

of about 1.1 seconds in speed in the one-finger work, a loss of about 1.6 seconds in the two-finger work, and a loss of about 1.3 seconds in the three-finger work. While these losses do not show a markedly greater decrement for the more heterogeneous tasks, they clearly indicate that in work of this general type the homogeneity-heterogeneity factor does not operate as it did in our earlier experiments. It looks very much as though the decrement produced in the one-finger work by virtue of its homogeneity were at least offset by the decrement which heterogeneity, presumably by increasing the competition factor, produces in the two- and three-finger work.

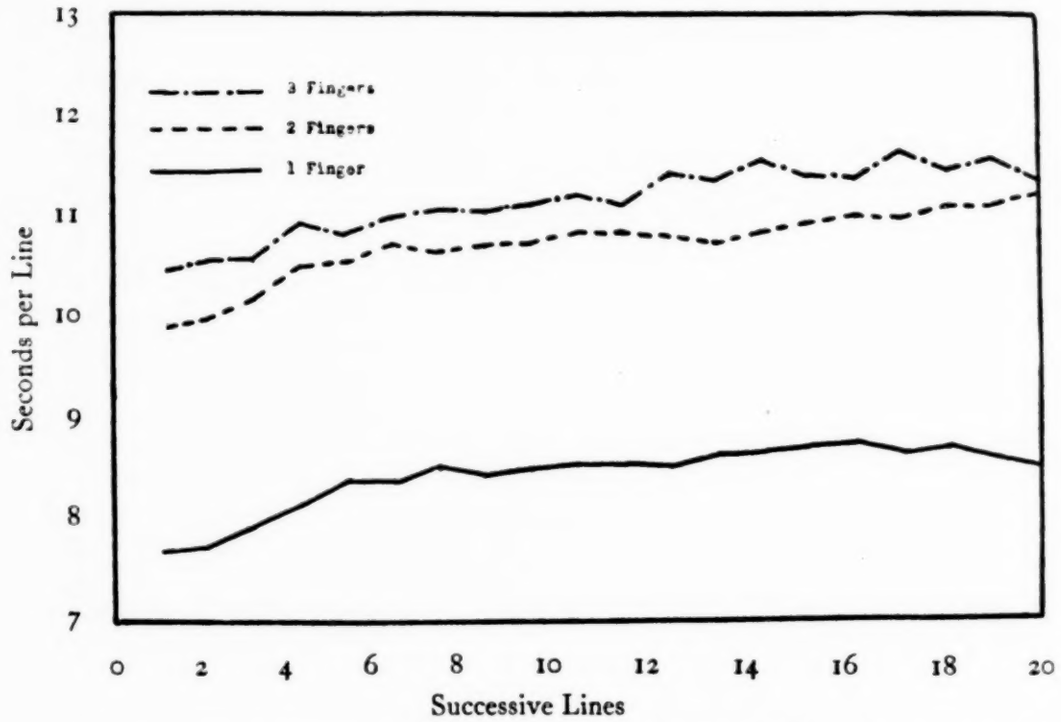

FIG. 14. Work-Curves for Experiment 4. Second Trials. See legend for Fig. 13.

In Fig. 14 we have the work-curves obtained from the second performances on each experiment-day. The results agree with those shown in Fig. 13. Certainly the one-finger work shows no more marked a decrement than do the tasks of greater complexity. For the single finger the total loss in speed is between 0.8 second and 1.1 seconds and for the combinations of fingers it is from 1.0 to 1.3 seconds.

The difference in general level of efficiency might suggest

the need of figuring the decrement in relative terms. Since, however, zero efficiency lies at some indeterminate speed and since that speed must be nearer that of the heterogeneous work than that of the more homogeneous work, it could add only something infinitesimal to the magnitude of the decrement in the former case.

Thus far, in all of the activities studied, we have been interested in decrements in speed of performance. As we have said, errors were negligible in the letter-writing and in the letter-naming. But in this last experiment there were many errors. This was especially true in the more complex tasks. These errors consisted of wrong letters struck and of failures to carry the key down to an actual impression. Now at the risk of elucidating the obvious we may call attention to the fact that where losses in accuracy are concerned, there is special likelihood of obtaining a more marked decrement in relatively heterogeneous work than in work of a homogeneous

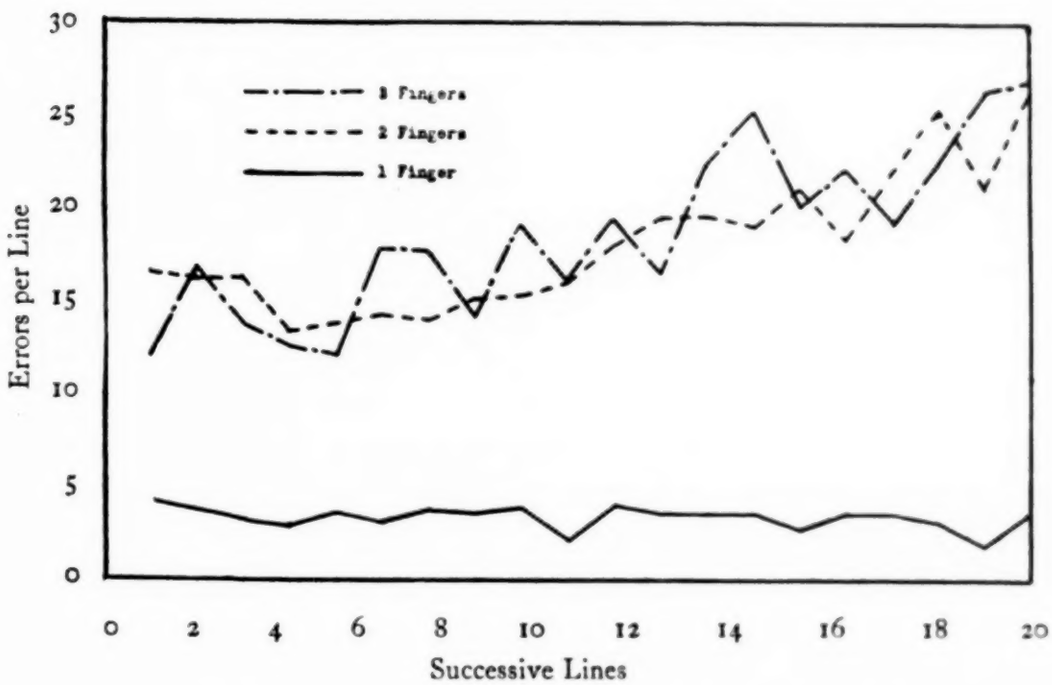

FIG. 15. Work-Curves for Experiment 4. First Trials. Horizontal axis represents successive lines written. Vertical axis represents total number of errors per line. Curves are totals for 21 subjects for 3 cycles.

character. In Fig. 15 we have work-curves in terms of errors for successive lines of typing. These curves are from the first

daily performances. Those from the second performances are so like those from the first performances that there is no reason to give them also. The curves representing two- and three-finger work show pronounced decrements, while none at all are shown by the curve for one-finger work.

Subjects' Reports: The typing experiment adds little, through subjects' reports, to what we have obtained. It was noticed that the typing proceeded more smoothly if not closely attended to. With the onset of pain in the hand and arm there were shifts in posture and in details of the movement which often brought relief. This is, of course, a very frequent occurrence in all forms of muscular work and one that makes the study of decrements in such acts somewhat difficult. If an act is complex enough to permit slight modifications without decreasing its productivity, production records may resist for a long time any serious decrement.

IV. Summary

In beginning an experimental study of factors underlying the variability of the work decrement we attempted an examination of the influence of the homogeneity-heterogeneity of the work itself. This factor we assumed to be important because of its very evident relationship to conditions of refractoriness. The large biological phenomenon of refractoriness is a direct correlate of the recency and frequency of functioning of the activity in which the refractoriness is manifest. Since relatively homogeneous work involves greater recency and frequency of functioning of each of its part-activities than does relatively heterogeneous work, we should expect psychological activities of a homogeneous sort to manifest more refractoriness and, therefore, more decrement than work of a more heterogeneous sort. The only problem was to discover activities in which variations in homogeneity-heterogeneity would be more important in determining decrement than other factors involved in those activities. In the rapid writing of simple alphabetical letter-groups and in the reading of letters printed at random on a card we found activities answering our requirements. In both cases we were able

to show a definite, simple relationship between homogeneity-heterogeneity and the magnitude of the work decrement.

We felt, however, that there might be many sets of tasks in which the variation of homogeneity-heterogeneity would not be the dominating factor in determining decrement. If the more heterogeneous of such tasks were also the more delicately adjusted, these more heterogeneous tasks might be as susceptible to the decrement effect as more heterogeneous tasks of the same general sort. The part-activities in the former case might tend to compete with each other and to offer great possibilities of interference, so that the fact that smaller elements of frequency and recency of functioning were present might be more than compensated for by the increased possibilities of competition. We sought an experimental verification of this conception. We were successful in finding the verification in certain movements involving the use of one, of two, or of three fingers. In the case of these movements there was something like an equal decrement for the tasks of all three degrees of heterogeneity. We believe that the greater refractoriness set up in the more homogeneous activity by virtue of a greater factor of repetition was more than offset as the acts became more heterogeneous, because that very increase in heterogeneity of composition allowed in this case for an increase in the amount of possible competition between part-activities.

We regard the experiments here reported as forming merely an entering wedge into the problem of the work decrement in psychological activities. The homogeneity-heterogeneity principle here demonstrated has, as we ourselves have shown, very definite limits. It may be obscured by a principle of competition. But we have here touched that principle only in the vaguest manner. One of us has pointed out [16] on general grounds the probable complexity of this principle and a thorough experimental study of it offers a real challenge. Other principles are also surely fundamental in the determination of the decrement and all of them offer rich territory for exploration.

[16] *Psychol. Rev.*, 1926, 33, 123–134.

READING REACTIONS FOR MATHEMATICAL FORMULÆ [1]

BY MILES A. TINKER

Psychological Laboratory, Stanford University

Review of Studies on Legibility

The purpose of this study was to investigate the relative legibility of exponents and subscripts in mathematical formulæ as compared to the same symbols when used in the body of such formulæ. During the last twenty-five years, the experimental study of reading has been very prominent; and current methods of teaching the art of reading, as well as the efficient construction of books, are now making considerable use of the results of these investigations. Since an appreciable amount of the reading of students involves formulæ which are found in mathematical and scientific texts, it would be instructive to know more about the processes concerned in these ocular functions. Also it is important that such reading be made as rapid and as easy as possible. Sanford (26) very aptly says: "The problem is to get the greatest amount of matter with the greatest ease in reading on the least space; or, as it has been phrased, to get the greatest legibility to the square inch."

Several investigations have concerned themselves with the legibility of the printed page. Some of the earliest of these experiments attempted to determine the legibility of the various letters of the alphabet and of various types of printing. As far back as 1881, Javal (19) came to the conclusion, after experimentation and observation, that the letters whose design was least desirable from the standpoint of legibility are: a c e f g r s t. Cattell (5, 6, 7) exposed both capital and small letters for very short intervals. In this way he determined the relative legibility of the isolated letters by taking the percentage of correct readings of the various letters

[1] A summary of this paper was given at the meeting of the Western Psychological Association, Mills College, California, June 18, 1926.

as exposed under these constant conditions. The order of capitals, from most to least legible, was found to be: W Z M D H K N X A Y O G L Q I S C T R P B V F U J E. For small letters the order was: d k m q h b p w u l j t v z r o f n a x y e i g c s. In a somewhat similar experiment with brief expositions, Sanford (26) used Snellen small letters. He found the order of the isolated letters, from most to least legible, to be: m w d q v y j p k f b l i g h r x t o u a n e s c z. Using the distance from the eye at which they became legible, he again determined the order of legibility of the same letters as: w m q p v y j f h r d g k b x l n u a t i z o c s e. Showing a group of nine letters at each brief exposure, Finzi (14) used the percentage of misreadings to compute the relative legibility of the letters. The descending order for legibility he found to be: P U A Q X T D S E W M V Y Z H C N F L R G B K O I. Finzi does not state what type or letter-form he used. Dockeray (9) obtained the legibility of isolated letters by determining the width of the field of distinct vision for each letter of the alphabet. The order of the letters from most to least legible for one of his three subjects was: m dk w p boqy jrx ag lvz istu hn c f e. The orders were similar, but not exactly the same, for the other two subjects. Shaw (27) summarized up to 1908 the results of the investigations on the legibility of small letters of the alphabet. The letters w m q p v j f he classed as most legible. Letters h r d g k b x l n u were classed as fair in legibility in comparison with the others; and the following letters were classed as poor: a t i z o c s e.

TABLE I

SUMMARY OF INVESTIGATIONS ON THE RELATIVE LEGIBILITY FOR THE LETTERS OF THE ALPHABET. THE ORDER IS (LEFT TO RIGHT) FROM MOST TO LEAST LEGIBLE

Finzi (1900).....	P U A Q X T D S E W M V Y Z H C N F L R G B K O I
Cattell (1886).....	W Z M D H K N X A Y O G L Q I S C T R P B V F U J E
	d k m q h b p w u l j t v z r o f n a x y e i g c s
Sanford (1888):	
Distance.....	m w d q v y j p k f b l i g h r x t o u a n e s c z
Short Exposure...	w m q p v y j f h r d g k b x l n u a t i z o c s e
Dockeray (1910)....	m d k* w p b o q y j r x a g l v z i s t u h n c f e

* Letters connected by an underscore were of equal legibility.

The outcome of the studies on the relative legibility of the letters has been brought together in Table I to facilitate comparison. The most legible letters begin at the left of the table. From inspection it is seen that, with the exception of Finzi, there is considerable agreement in the results. In several instances a given letter occupies the identical rank in two or more investigations. In Sanford's two methods of procedure, e.g., v, y and j were found in the same rank. The letter m gained the same position by Sanford (distance) and Dockeray, and t and e the same by Sanford (short exposure) and Dockeray. Similar identical placements appear in comparing Cattell's results with those of the others. And where letters occupy different relative positions in the various studies, many of them are still in the same quartile. The amount of correspondence is best shown by correlating the different rankings with each other. This has been done by the rank-order method in Table II. Finzi's results do not

TABLE II

CORRESPONDENCE AMONG THE INVESTIGATIONS ON RELATIVE LEGIBILITY OF LETTERS AS SHOWN BY RANK-ORDER CORRELATION

	Cattell: Capitals		Cattell: Small Letters		Sanford: Distance		Sanford: Short Exposure		Dockeray	
	ρ	P.E.$_\rho$	ρ	P.E.$_\rho$	ρ	P.E.$_\rho$	ρ	P.E.$_\rho$	ρ	P.E.$_\rho$
Finzi	−.02	±.14	—	—	—	—	—	—	—	—
Cattell:										
Capitals			.24	±.13	.14	±.14	.15	±.11	.38	±.12
Small Letters					.62	±.09	.57	±.09	.58	±.09
Sanford:										
Distance							.88	±.03	.68	±.08
Short Exposure									.64	±.08

correspond at all with Cattell's capitals. It is more than possible that he used a type or letter-form which prevents a reliable comparison with the other studies. The positions of Cattell's capitals show little correspondence with the positions of small letters in any of the experiments. The highest correlation coefficient here is only .38 ± .12. This is to be expected because of the difference in form of capitals and small

letters. The various placements of the small letters correspond fairly well, the coefficients ranging from .57 ± .09 to .88 ± .03. It is quite likely that these orders of legibility will not be materially altered by future investigations so long as the letter-forms remain about as at present.

But there have also been studies so designed as to determine the relative legibility of the same letters or words when printed in different styles of type or when written by the hand. The effect of size, face, ornateness of type, as well as spacing, have been so studied. In an experiment to determine the influence of size and form on the legibility of letters, Griffing and Franz (17) found that it was always easier to read larger types than smaller; that, as the size of type decreased, the fatigue from reading rapidly increased; and that one could always read Gothic letters more easily than Roman. Roethlein (25) investigated the relative legibility of 26 different faces of printed types. The letters were presented singly and in groups. She came to the conclusion that legibility is a product of six factors; *viz.*, the form and size of the letter, the heaviness of the letter-face (the thickness of the lines which constitute the letter), the width of the white margin which surrounds the letter, the position of the letter in the letter-group, and the shape and size of adjacent letters. In working out a more legible form for a telephone directory, Baird (1) found that increasing the leading between lines by a small amount was an important aid to legibility. Huey (18), Blackhurst (4), and Bentley (3) have also shown the importance of leading in the legibility of the printed page. After reviewing several studies, Gray (16) concludes: "(1) Legibility varies, to a greater or less extent, with style of type. The differences are much less when letters are presented in groups than when letters are presented in isolation. (2) Plain styles of type, such as Roman and Caslon, are much more legible than the more ornate styles of type, such as Old English and German type. . . . (3) Handwritten and typewritten materials are less legible than material printed in plain type."

Another factor in legibility enters when combinations of

words are printed in lines for reading material. It is generally agreed that very long lines of print tend to lessen legibility. After experimenting, Dearborn (8) came to the conclusion that a line from 75 to 85 mm long and with type 1.5 mm high gives a satisfactory and efficient reading line. Huey (18) thought that 90 mm should be the maximal line-length and 60 to 80 mm the optimal. Both Dearborn (8) and Huey (18) agree that regularity in length of line is an important factor in reading, because of the 'short-lived motor habits' set up by the eye in reading lines of uniform length. In an analytical study Judd (20) found that doubling the size of eleven-point type, or reducing it to half-size made only a slight difference in efficiency for adult reading. In a more elaborate study Gilliland (15) recorded eye-movements, using two series of reading material which ranged in size of type from 6 to 54 points and from 3 to 90 points. He concluded that "the reading of the average adult is not greatly affected by changes in the size of type between the limits of 36-point and 6-point type." In going over the various investigations, Gray (16) reports that it is generally agreed that printed letters should be a deep black, and the paper pure white and without gloss. The influence of colored paper on ocular efficiency has not been thoroughly investigated and no definite conclusion can be made concerning it. In England, the Industrial Fatigue Research Board is carrying out an investigation (13) to determine the most effective contrast between colored papers and inks. Preliminary results suggest that color contrast is of comparatively small importance in hygienic printing. Starch (28) reports a study which agrees with this view.

The Plan of Our Investigation with Formulæ

In this investigation, which is the preliminary part of a larger study on the reading of formulæ, the attempt has been made to compare the legibility of certain parts of mathematical formulæ by comparing the reading-time of the digits 1 to 5, and the letters n, m, x, y, and t, when occurring in the body of the formulæ, to the reading-time when they occur as exponents or subscripts. This comparison is designed to throw some

light on the relative legibility of the parts considered when the reading is done under approximately normal reading conditions. Most of the previous experiments which have investigated the legibility of isolated letters, words, or like material, have used some form of short-exposure apparatus. While this kind of experiment has achieved noteworthy results, it is desirable here to provide a closer approximation to the normal reading situation. In his outstanding study of visual fixation Dodge (10) gave some consideration to this and later when investigating the psychological effects of alcohol (11) he devised a technique and apparatus for the determination of the reading-time of four-letter English words. In this latter publication he states the following important conditions. The stimulus should appear suddenly, all at once, in the field of clear vision, in order to increase to a maximum the visual controls which ordinarily complete the process which is begun in the prefixational perception of the reading material. There should be provision for satisfactory adaptation to distance and to illumination. The stimulus should remain in view until the act of perception has been completed and the response made. Continuation of the exposure after perception has taken place may be used as a check on misperception. The fulfilment of these conditions seemed to us to offer the best possibilities in our experiment.

Three faculty members and four students in the Department of Psychology at Stanford were subjects in this experiment. Five of these had the equivalent of about two courses in high school mathematics, and the other two had completed calculus. This arrangement gave us a mediocre and a skilled group so far as mathematical training goes. All were tested with the Snellen chart for visual acuity. Those who wore glasses were allowed to keep them on in the visual test since they wore them while observing in the experiment. They all tested 20/20, or above, on the chart.

In order that the material to be read might approach as closely as possible the real conditions encountered by the student, mathematical formulæ were cut from 4 standard algebra texts (12, 23, 24, 30) and pasted on small cards about

7 cm by 3 cm in size. These formulæ were of various sorts and lengths. About ½-inch below the left end of each formula a large black dot was placed for pre-exposure fixation. The distance from the left end of the formulæ to the parts to be read could not be exactly controlled. On the average, the exponents and subscripts were 0.5 of a letter-space nearer the left end than the digits, and 0.3 of a letter-space nearer the left end than the letters, in the body of the formulæ. Therefore any advantage derived from being nearer the left end of the formula would go to the exponents and subscripts.

Apparatus for Reading Reaction [2]

The apparatus used for exposing the formulæ and recording responses was the same general scheme as described by Dodge (11), but with the modifications introduced by Miles (2, 22) for freeing the experimental procedure from several distracting factors. Its main parts are: the exposure apparatus, a voice key, two electromagnetic markers, a large time-clock, and a Baltzar kymograph. The exposure apparatus was on a wooden screen large enough to hide from the observer the kymograph and markers, as well as the manipulations of the experimenter. In this screen, at a level with the observer's eyes, was a small window about 2 cm by 7 cm in extent. Behind this window was a light horizontal arm with one end attached to the screen and the other free to move past the window. The free end carried a pack of from 12 to 15 stimulus cards. The card-holding device was improved by Miles to hold several cards (he used packs of 25) instead of one as used by Dodge. With this arrangement the experimenter was able to slip out one of these cards after each exposure without any interruption of the kymographic record. Only the black dot on the stimulus card was visible to the subject during the pre-exposure interval, as the card-carrying arm was held up by an electro-magnet. This magnet was in circuit with the voice key, a marker and a switch on the kymograph. At

[2] This equipment was largely made possible by the Thomas Welton Stanford Fund for Psychological Research. The apparatus, as shown in Fig. I, is mounted on a small table and is to be used as a permanent reading-reaction unit. It is economical in space occupied and can be easily moved as a unit from room to room.

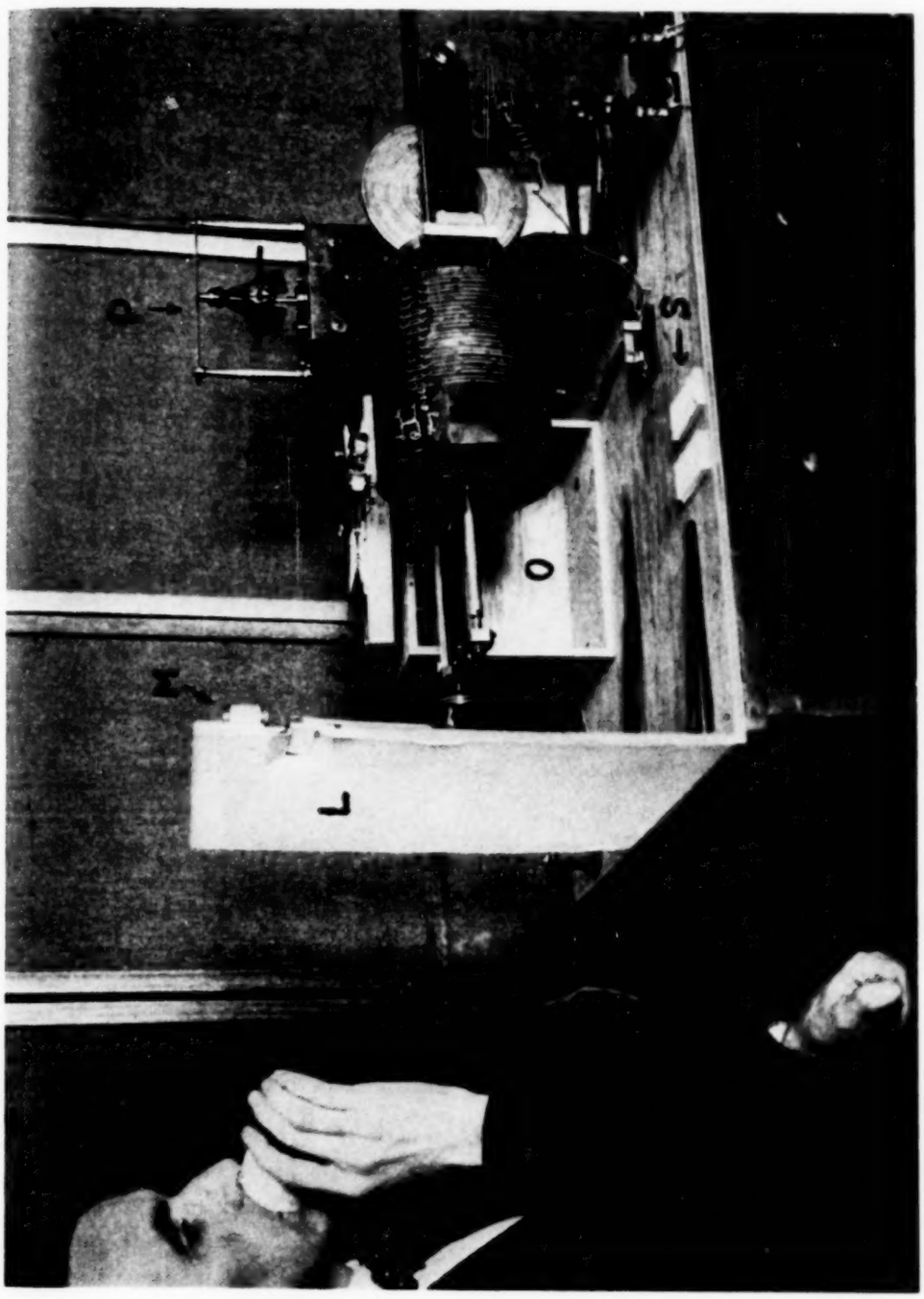

FIG. 1. Apparatus and conditions for reading reactions. *K*, voice key; *L*, screen with windows; *M*, magnet for holding up arm of exposure device; *N*, markers attached to mounting board which is hinged to box-support *O* (This board is turned up and back, carrying the markers away when the drum is changed); *P*, kymograph revolving drum with spiral motion; *R*, rheostats; *S*, groups of stimulus cards.

each half revolution of the kymograph drum, an automatic circuit breaker attached to the shaft of the kymograph momentarily broke this circuit, releasing the free end of the arm carrying the cards, and at the same instant it caused the marker to record on the smoked kymographic paper. The arm quickly dropped far enough to bring the formula to the middle of the window at the spot where the fixation dot had been exposed. This arm was activated by a stout rubber band and was stopped suddenly and held firmly in the center of the window with very little noise and no vibration, by means of what Dodge calls the 'pendulum stop.' The main feature of this stop consists of two bars, one short and one long, riveted together in a movable joint. The long end of the lever thus made was attached to the movable end of the exposure-arm. The short end was attached to the screen directly below the attachment to the arm. When the exposure-arm was up the short part of the lever was horizontal, and when down, vertical, and thus parallel to the long bar of the lever.[3] In this position, the arm holding the stimulus could move neither up nor down until the experimenter pushed the joint of the lever to one side. The stimulus card remained exposed until the experimenter reset the apparatus for the next exposure. With this type of exposure, the suddenly dropping card came to a rest without any rebound or secondary vibration so that the clearing up of the visual field was wholly physiological. Thus one obtained an exposure that has a minimum of interference with 'the normal and adequate perception' of the stimulus. The subject reacted by speaking into the voice key. This broke the circuit again and thus the marker recorded a second time on the same line as contained the exposure mark. A time-line giving the time in seconds ran parallel to the reaction-line. This made it easy to read the time interval between the beginning of the exposure and the beginning of the response. The subject sat with his eyes at a level with the exposure-window and at a comfortable reading distance from the stimulus card. With the exception

[3] For Dodge's description see Dodge and Benedict, Psychological effect of alcohol, *Carnegie Instit. Washington*, 1915, no. 232, p. 94, fig. 13.

of the first few observations these distances were measured at each sitting. They varied with the subjects from 13 to 16 inches. An attempt was made to have each subject keep approximately the same distance in successive sittings. Fifty

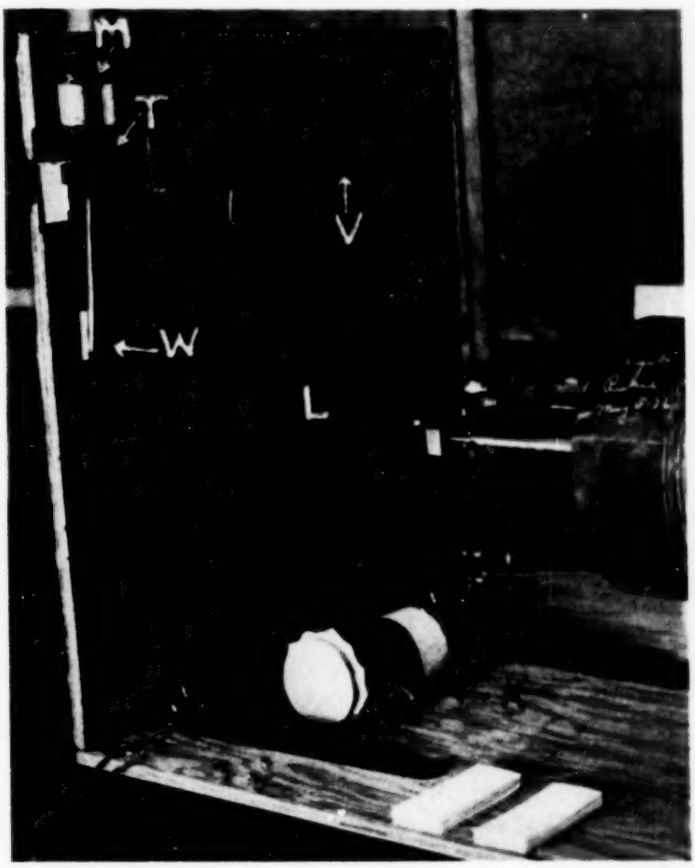

FIG. 2. Detailed view of the Dodge drop-exposure. *K*, voice key with sanitary paper protector for applying to face; *L*, rear of screen; *M*, magnet; *V*, card-carrying arm; *T*, card pack-holder; *W*, double pendulum-stop.

exposures could be run off in about 20 minutes. This exposure apparatus is easily and inexpensively constructed. It should be found useful in many experiments involving visual perception.

The voice key consists of a short brass tube about $1\frac{7}{8}$ inches in diameter, one end of which is formed to fit against the face surrounding the lips. Before using, a clean piece of paper is fitted around the speaking end to insure sanitation and also to protect the skin from contact with the cold metal. Into

the other end is fitted a hard rubber ring which holds a spring break-key. A thin rubber membrane is stretched across this end of the cylinder so that it holds the spring key closed. Air pressure, caused by speaking into the cylinder, pushes out the membrane and allows the key to open. A vent in the side of the cylinder allows the air to escape and the membrane to readjust itself.

The stimulus marker and the time marker were mounted, side by side, on a board 1 x 6¼ x 10 in., hinged to a wooden box support.[4] When not in use or when a drum was to be changed the board could be turned back on its hinges and away from the kymograph. This practical arrangement made the adjustment of the markers for successive records both easy and accurate. A small rheostat was inserted in the time-circuit and the exposure-response circuit.

Experimental Procedure

The experiment was conducted under artificial illumination which was (as determined by the Foot-Candle Meter of the Edison Lamp Works) 14 foot-candles per square foot. (It is recommended that a drafting room have an illumination of 15 foot-candles per square foot and a school room 8 foot-candles.)

Each subject observed at regular periods for 9 weeks. The first week was devoted to practice. In the fore part of each week for the first 7 weeks, the subjects read digits and letters in the body of the formulæ and, in the latter part, exponents and subscripts. On one-half of the days that digits and letters were read, digits were read first, and on the other days letters were read first. At the beginning of the experiment, the subjects were told the purpose of the study and were instructed how to use the voice key. Each day before reading digits the subjects familiarized themselves with the following instructions:

> When the ready signal is given look at the dot on the card exposed in the window. Shortly after the ready signal, the card will suddenly drop down enough to expose a

[4] This feature of technique for adjusting markers to smoked drums has been developed by Professor Miles at the Stanford Laboratory. The author recommends it as having several advantages over ring-stand supports, especially in semi-permanent setups. Not the least advantage is that the markers are held without vibration.

formula. Do not attempt to anticipate or move the eye in the direction of the formula, but wait until it appears. Speak the first digit you come to in the body of the formula, reading as usual from left to right, and reading numerators before denominators where fractions are involved. Exponents and subscripts do not count in this experiment and are not to be pronounced. Do this as quickly as you can but be sure that you are correct.

Before reading letters or subscripts the subjects familiarized themselves with instructions similar to the above but appropriately modified. About $1\frac{1}{2}$ seconds before each exposure the experimenter gave the ready signal.

At the end of each sitting the subject was asked to give a written report discussing the facility or difficulty he encountered in making the observations, the place where difficulties occurred, and his degree of assurance.

For the observations during the second week an entirely new set of formulæ was used. The formulæ used during the first week were employed again in the third week, but in a different order, which was determined by chance. This arrangement was followed throughout the experiment.

During the eighth week the same formulæ were used as during the previous weeks but the procedure was changed. The subject reacted by speaking the word 'found' into the voice key as soon as he came to the first digit. A similar routine was followed for letters and for exponents and subscripts.

Using the same apparatus and the method employed during the first seven weeks, the average reading-time for each subject was found for fifty four-letter, monosyllabic English words. This material made it possible to compare the 'formula' reactions with reactions for words, the reading of which involved no selection from the material presented on the screen. In order to familarize the subjects with these words they were arranged in an order different from that employed in the experiment, and were read over once before the experiment began.

A series of typewritten formulæ was constructed, in which the exponents were considerably larger than those found in the algebras, although smaller than the typed characters in the body of the formulæ. The reading-times of the two sym-

bols *n* and *2* were used from the data collected in the reading of these typed formulæ. The formulæ containing these two characters were arranged in chance order among the whole set of typewritten formulæ read. The distance of *n* and *2* from the left end of the formulæ was, on the average, just the same when used as exponents as when employed in the body of the formula. The same experimental procedure was carried out as in the first experiment.

Experimental Results

The intention of this experiment was not to obtain absolute values for reading parts of mathematical formulæ, but rather to obtain these reading-reactions under constant conditions for the purpose of comparison. One of the most striking things demonstrated by objective studies of reading is the great range of individual differences in speed and comprehension. The individual differences among our subjects in reading parts of mathematical formulæ were also pronounced. Since they were as large as the difference between the naive and the mathematically skilled observers, and since all the records showed the same tendency, all the results have been averaged together. In considering these results one must keep in mind the number of subjects employed. The conclusions arrived at are tentative and suggestive rather than final.

In Table III are shown the average reading-reactions for symbols read as exponents or subscripts, and also symbols read in the body of the formulæ by seven subjects. Exponents and subscripts were put in the same category in this experiment for two reasons: (1) They are both small and of the same size; and (2) they are both out of alignment with the horizontal line of reading. The written observations of the subjects seem to have justified this procedure. Some of the subjects claimed that the exponents were harder than subscripts to read, a few said that the subscripts were harder, and still others stated that the two were of equal difficulty. Some of the observers were not consistent as they changed their opinion in this regard at different sittings.

It is seen from Table III that, on the average, letters were

read .04 sec quicker than digits when both appeared in the body of formulæ. Nevertheless every subject reported that letters in this position were harder to read than digits. The

TABLE III

READING REACTIONS FOR CERTAIN PARTS OF MATHEMATICAL FORMULÆ
(TIME IN 1/100 SEC.)

Position within Formula	Kind of Reaction	Digits and Letters Used											
		1 or N		2 or M		3 or X		4 or Y		5 or T		Average	
		Av.	A.D.	Av.	A.D.	Av.	A.D.	Av.	A.D.	Av.	A.D.	Av.	A.D.
Exponent or subscript	Digit	68	14	84	18	76	17	78	11	73	11	74	14
	Letter	83	21	87	24	80	18	76	19	93	28	83	22
In body of formula	Digit	76	21	67	16	60	12	70	14	68	17	68	16
	Letter	67	14	61	11	65	11	62	13	66	17	64	13

difference is not large enough to be of much significance but it is interesting because when these same letters and digits are read as exponents or subscripts a more significant difference does appear in the opposite direction. There the average reading-time of digits was .09 sec faster than for letters. This greater speed for digits was largely due to the rates found for reading the symbols *1* and *5*. The longest reading-time was for *t* when read as an exponent and as a subscript.

TABLE IV

DIFFERENCES IN READING REACTIONS FOR CERTAIN PARTS OF MATHEMATICAL FORMULÆ. EXPONENT AND SUBSCRIPT TIME MINUS BODY TIME.
(TIME IN 1/100 SEC.)

Reading Difference for:	1 or N		2 or M		3 or X		4 or Y		5 or T		Average	
	Av.	A.D.	Av.	A.D.	Av.	A.D.	Av.	A.D.	Av.	A.D.	Av.	A.D.
Digits	−8	−7	17	2	16	5	8	−3	5	−6	6	−2
Letters	16	7	26	13	15	7	14	6	27	11	19	9

Comparison of the differences in reading-times of symbols as exponents and subscripts and symbols in the body of the

formulæ is shown in Table IV. It is seen that the digit *1* was read nearly 1/10 sec faster when a subscript than when in the body of the formula. Inspection of the written observations and of the formulæ employed as stimuli gives a possible explanation of this. In the body of the formulæ *1* nearly always occurred as part of a number such as 12, 10, or 130. Every subject reported the tendency to speak the whole number and this had to be inhibited before *1* could be pronounced. As subscripts, the *1*'s were nearly always discrete. Most of the subjects observed that discrete symbols were comparatively easy to read. This agrees with Roethlein (25) who found that the legibility of letters was greater when they were presented in isolation than when presented in groups. Digits *2* and *3* were read appreciably faster in the body of the formula, while for *4* and *5* there is not such a significant difference. The fact that the digits *1*, *4* and *5* have more individuality in form, especially the upper half of *4* and *5*, indicates that probably these digits are more legible than the others. Whipple (31) says that, in digit cancelation tests, *0* and *4* are harder to cancel than *1* or *7*, and that *6* and *9* are the most difficult of all. Messmer (21) claims that form is related to perception. He states that breadth, height and geometrical form are significant characteristics of letters. According to Shaw (27) and Huey (18) the upper half of letters and words in the printed line is of greater importance for correct perception than the lower half. The variation in reading-times for letters as exponents or subscripts was appreciably greater than for letters read in the body of formulæ, while that for digits showed little difference. With letters there is a consistent trend. Every letter in the body of formulæ is read significantly faster than the corresponding exponent or subscript. The range of differences for reading in these two positions was .14 to .27 sec, which was for *y* and *t* respectively. The average difference between exponent or subscript versus body was only .06 sec for digits but was .19 sec for letters. This shows rather emphatically that the five letters lose comparatively more in legibility when used as exponents or subscripts than do the digits.

TABLE V

AVERAGE READING REACTIONS FOR VARIOUS TYPES OF MATHEMATICAL MATERIAL AND FOR 4-LETTER ENGLISH WORDS
(TIME IN 1/100 SEC.)

Kind of Reaction	Exponents and Subscripts		Digits		Letters	
	Av.	A.D.	Av.	A.D.	Av.	A.D.
'Found'	61	11	55	10	53	9
Naming when 'Found'	97	13	88	11	86	8
Typed Formulæ	57*	8	53	13	—	—
	67†	9	—	—	63	11
Text-Book Formulæ	79	18	68	16	64	13
Naming 4-letter words	42	3				

* This horizontal line represents reaction to the digit **2**.

† The letter *n* was used as stimulus in this line of reactions.

Table V gives the average reading-times for various kinds of mathematical material and for 4-letter English words. In the first row are the reaction times when the subjects spoke 'found' as soon as they had discovered the position of the desired symbol without waiting to perceive its identity. This experiment was performed with the hope that it might throw some light on the influence of position on reading-times for exponents and subscripts. Every subject reacted with the word 'found' in a shorter time than he could name the corresponding character. The first column of Table VI shows that the averages of these differences are significantly large. In Table V are given the average reading-times for naming the different items after they had been found. They maintain the same relative position as the reading-times determined in the first part of the experiment. Naming the symbol after saying 'found' was merely for the purpose of checking the correctness of responses. The same table indicates that it takes a little longer to find exponents and subscripts than it does to find the same characters in the body of formulæ. These differences are about the same as those between reading digits in the two positions, but much less than the difference between reading letters in the two positions. In the written reports, most of the observers claimed that they knew what a

letter or digit was as soon as it had been found, but did not know the name of exponents or subscripts as soon as they had localized them. This small difference in time between 'finding' exponents or subscripts and characters in the body of formulæ together with the reports of the observers suggests that position is one of the factors which make exponents and subscripts less legible, but is not the only factor.

Table VI

Differences in Reading Reactions for Parts of Mathematical Formulæ in Various Types of Material (Time in 1/100 Sec.)

Parts	Text-Book Formulæ Minus 'Found'		Naming After 'Found' Minus Text-Book		'Found': Exponent and Subscript Minus Body		'Typed': Exponent Minus Body	
	Av.	A.D.	Av.	A.D.	Av.	A.D.	Av.	A.D.
Exponents and Subscripts	18	7	19	6	—	—	—	—
Digits	13	6	20	−5	6	1	4	−2
Letters	11	4	22	−5	8	2	4	−5

The experiment in which typed formulæ were read was intended to throw further light on the influence of position on legibility of exponents and subscripts, as well as the influence of size. The symbols *n* and *2* were chosen to be reacted to because they had been read with significantly less times in the body of formulæ in the first experiment. In these formula the exponents were large when compared to those cut from text-books, but were smaller than the symbols in the typed body of the formula. The exponents were always discrete. Any difference in reading-times for the different parts of these formulæ would probably be due, in part at least, to the position of the exponents. The reading-times are found in Table V. Table VI shows that both *2* and *n* were read in the body of the formulæ .04 sec faster than as exponents. From this and from the results of reacting with 'found' it is indicated that position is one of the factors which lessen legibility of exponents or subscripts when compared with the characters in the body of formulæ. The shorter reading-times when compared with text-book formulæ indicate that size is another factor of legibility.

At the bottom of Table V is shown the average reading-time for familiar four-letter English words. The reading time for the seven subjects was .42 sec with an average deviation of .03 sec. In a similar experiment with four-letter English words, Dodge (11) found an average reading time of .46 sec with an average deviation of .04 sec for seven subjects. The shorter time for the subjects in this experiment is very likely due to the large amount of experience they had had with our apparatus. It should be noted that this reaction-time, which involved no selection from material exposed, is considerably shorter than any of the reading-times for parts of formulæ. The reading of only a single designated part of the material exposed does, to some degree, set up an artificial reading situation. Throughout this experiment, however, this artificial reading situation has been held constant for the things compared and therefore it should not invalidate the results.

Analysis of the Reports

The written reports of the subjects show some of the difficulties encountered in reading parts of mathematical formulæ and also throw some light on how formulæ are read. An analysis of the reports indicate that there were six different sorts of difficulty involved in reading certain parts of formulæ in this experiment.

(1) *The Kind of Thing to Be Read.*—Letters in the body of formulæ were considered a little harder to read than digits. Nevertheless Table III shows that letters in this position were read with somewhat shorter times than digits. Exponents and subscripts were said to be much more difficult to read than the corresponding symbols in the formula body. The fact that they are scattered over a wider vertical range was considered as one of the factors here involved. This checks the objective data. A greater effort of attention was required when reading exponents and subscripts. Letters as exponents or subscripts were considered more difficult to read than digits in the same position. This also agrees with the objective results. For some strange reason letters are always smaller and often less distinct in heaviness of face than the digits

which are printed with them, especially when used as exponents or subscripts. This adds to the poor legibility inherent in certain letters. As noted before, exponents and subscripts were considered to be of equal difficulty.

(2) *Certain Letters.*—In the body of formulæ *t* gave the most trouble. The letters *x* and *y* were called difficult and *m* and *n* easy in this position. The digit *4* was perceived with great facility in the body. As exponents or subscripts, the letters *x*, *y*, and *t* were considered much harder to read than *m* and *n*. One observer gave the order of facility from easy to hard as *m*, *n*, *t*, *y*, *x*. By referring to Table I, it can be seen that those letters which gave the most difficulty are usually classed as poor in legibility. In spite of the outcome of the studies on the legibility of letters, the writers of mathematical texts persist in using letters which are less desirable from the standpoint of reading.

(3) *Size.*—The small size of digits and letters as they occurred in exponents and subscripts was considered a source of difficulty. The digits and letters presented unequal difficulties, however, in these positions. Letters were harder to read because they seemed smaller, less definite and more complicated in structure than digits. The subjects stated that the *x* as read here seemed particularly small and indistinct. One observer stated that the size of the letter or digit and its position in the formula were of more importance than the particular symbol read. Size as a source of difficulty agrees with the objective results of the experiment and also with the results of other investigators (25, 17).

(4) *Fractions.*—The need of shifting considerably the alignment of the eyes where fractions were involved offered obstacles to easy reading. In most subjects there was a marked tendency to continue with the eye right along the horizontal line of print rather than to shift the eye down to read the denominator, when one was present. This indicates, probably, that the reading of mathematical material which involves a large amount of fractions is a difficult and fatiguing task, except possibly for those who are habitual users of such subject matter.

(5) *Complexity.*—A relatively complex formula consciously hindered the ready perception of the desired character. Long formulæ were more difficult than short. Confusion tended to enter even though the desired symbol was at the beginning of the formula. A discrete, uncomplicated exponent or subscript was comparatively easy to read while two or more exponents or subscripts in close spatial relation set up a tendency to read the whole group or to pronounce the most conspicuous member of the group. This agrees with Roethlein's (25) results on the legibility of letter-groups. The presence of radical signs, brackets and unusual constructions in formulæ caused complexity and thus harder reading. When the character to be read was at the right end of a long formula or to the right of an equality sign, a delay in response was often noted.

(6) *Configuration.*—A combination of digits or of a digit and a letter was often perceived as a whole and the subject had to inhibit the tendency to speak the whole combination before he could read the desired symbol. When such numbers as 12, 130, or 386 occurred and the subject was to read the first digit, there was a decided tendency to speak the whole number. In an eye-movement study, Terry (29) found that certain familiar numerals seemed to be read as wholes. Some subjects read two digits at a single fixation, some three and one as many as four. This may indicate that where there are numbers of different lengths, there is some grouping of digits into perceptual units. In our later study on the reading of mathematical formulæ this subject will be given further consideration.

Some of the subjects were aware of eye-movement in the reading of parts of formulæ. The saccadic movements in the horizontal were the type noticed. These movements became especially noticeable where the formulæ were long and the desired letter or digit was well toward the right end.

The assurance that the responses were correct was always high. Assurance here, of course, can only refer to the interval between the exposure and the beginning of the vocal reaction. While he is speaking, the subject may notice that he is in error.

Summary and Conclusions

1. Investigations of various authors on the relative legibility of the small letters of the alphabet agree fairly well with each other in results obtained. Some letters are designated as high in legibility, others as only moderate, and still others as low. In spite of these discoveries, letters such as *a*, *t*, *i*, *z*, *c*, *s*, and *e*, which are low in legibility, and *n*, *x*, *k* and *r*, which are only fair in legibility, are widely used in formulæ, many of them in very small type, as exponents and subscripts. The more legible letters should be made use of in place of the less legible in the construction of mathematical and related formulæ.

2. In the body of formulæ, letters were read slightly quicker (.04 second) than digits. As exponents or subscripts, however, letters were read .09 second slower than digits. This indicates that the legibility of these letters decreases with decrease in their size at a greater rate than it does for the digits.

3. The average increase in reading-time of digits as exponents and subscripts over that in the body of formulæ is .06 second. The digits *2* and *3* were the only ones that had really significant increases. This suggests that *2* and *3* are less legible than the other digits used.

4. On the average, letters as exponents and subscripts were read .19 second slower than in the body of formulæ. Every letter showed a significant trend in this direction. Therefore these letters, as printed in standard mathematical texts, are considerably less legible as exponents and subscripts than in the body of formulæ. These letters lose comparatively more in legibility than do the digits when both are used as exponents and subscripts.

5. Digits and letters in any part of formulæ could be 'found' appreciably quicker than they could be named. The identity of any digit or letter in the body of a formula was perceived as soon as it was found. As exponents or subscripts, however, the identity of the symbols was not perceived as soon as they were localized. Digits could be 'found' .06 second quicker, and letters .08 second quicker than exponents and subscripts. This indicates that position is one of the

factors in the legibility of exponents and subscripts, but not the only factor.

6. In typed formulæ, where exponents were discrete and the factor of size materially lessened, the letter and digit were read .04 second faster in the main body than as exponents. When text-book formulæ were used, these differences for *n* and *2* were .16 and .17 second respectively. This indicates that size is a factor in the legibility of exponents and subscripts. It also supports the contention that position is also a factor.

7. The average reading-time for four-letter English words was .42 second, which is considerably shorter than the times for any of the formulæ read. This shows that reading which does not involve selection from material presented is easier than that which does.

8. The difficulties encountered in reading certain parts of mathematical formulæ were of six different kinds: (1) Letters were considered a little harder than digits in the formula-body and the letters were called much harder than digits used as exponents and subscripts. Exponents and subscripts were termed much more difficult than the symbols in the body. (2) In the body of the formulæ *m* and *n* were considered easy, *x* and *y* more difficult, and *t* very difficult to read. As exponents and subscripts, though all letters were hard, the same rank-order of difficulty held as in the body. (3) The small size of digits and letters affected the legibility of letters more than that of digits. (4) The need of shifting considerably the alignment of the eyes where fractions were involved offered obstacles to easy reading. (5) A relatively complex formula consciously hindered easy reading. Factors influencing complexity were; long formulæ, compound exponent or subscript, radical signs and brackets. (6) The tendency to perceive combinations of digits or a digit and a letter had to be inhibited before reading the single item.

9. The subjects were all quite confident that their responses were correct.

10. There seems to have been no intelligent effort to control to the best advantage the size, heaviness of face, and the width of the white margin surrounding the symbol in regard to the legibility of the printed formula.

11. An effort should be made to arrive at a more legible and yet economical printed page of mathematical formulæ.

References

1. Baird, J. W. Legibility of a telephone directory. *J. Appl. Psychol.*, 1917, **1**, 30–37.
2. Benedict, Miles, Roth & Smith. Human vitality and efficiency under prolonged restricted diet. Carnegie Institution of Wash. Publ., no. 280, 1919, 165–167, 602–605, and Fig. 31.
3. Bentley, M. Leading and legibility. *Psychol. Monog.*, 1921, **30**, no. 136, 48–61.
4. Blackhurst, J. H. Leading as related to readability in first four grades. *School & Society*, 1922, **16**, 697–700.
5. Cattell, J. McK. The inertia of the eye and brain. *Brain*, 1886, **8**, 295–313.
6. —— Ueber die Zeit der Erkennung and Benennung von Schriftzeichen, Bildern und Farben. *Phil. Stud.*, 1885, **2**, 634–650.
7. —— Psychometrische Untersuchungen. *Phil. Stud.*, 1886, **3**, 305–335; 452–493.
8. Dearborn, W. F. The psychology of reading. New York, 1906; also as *Columbia Univ. Contrib. Phil. and Psychol.*, 1906, **14**, no. 1.
9. Dockeray, F. C. The span of vision in reading, and the legibility of letters. *J. Educ. Psychol.*, 1910, **1**, 123–131.
10. Dodge, R. An experimental study of visual fixation. *Psychol. Rev. Monog. Suppl.*, 1907, **8**, no. 35, p. 54, Fig. 3.
11. Dodge, R., and Benedict, F. G. Psychological effects of alcohol. Carnegie Institute of Wash. Publ. no. 232, 1915, 90–108.
12. Egerton, E. I., and Carpenter, P. A. A first course in Algebra. New York, 1923.
13. Fifth Annual Report of the Industrial Fatigue Research Board, p. 15. H. M. Stationery Office, London, 1925.
14. Finzi, J. Zur Untersuchung der Auffassungs- und Merkfähigkeit. *Psychol. Arbeiten (Kraepelin)*, 1900, **3**, 289–384.
15. Gilliland, A. R. The effect on reading of changes in the size of type. *Elem. School. J.*, 1923, **24**, 138–146.
16. Gray, W. S. Summary of investigations relating to reading. *Suppl. Educ. Monog.*, 1925, no. 28.
17. Griffing, H., and Franz, S. I. On the conditions of fatigue in reading. *Psychol. Rev.*, 1896, **3**, 513–520.
18. Huey, E. B. The psychology and pedagogy of reading. New York, 1908.
19. Javal, E. L'évolution de la typographie. *Révue scientifique*, 1881, **26**, 802–814.
20. Judd, C. H., and others. Reading: its nature and development. *Suppl. Educ. Monog.*, 1918, no. 4, 15–53.
21. Messmer, O. Zur Psychologie des Lesens bei Kinder und Erwachsenen. *Arch. f. d. ges. Psychol.*, 1903, **2**, 190–298.
22. Miles, W. R. Alcohol and human efficiency. Carnegie Institution of Wash. Publ., no. 333, 1924, 87, 89.
23. Milne, W. J., and Downey, W. F. First year algebra. New York, 1924.
24. Rietz, H. L., and Crathorne, A. R. College algebra. New York, 1924.
25. Roethlein, B. E. The relative legibility of different faces of printing types. *Amer. J. Psychol.*, 1912, **23**, 529–546.

26. SANFORD, E. C. The relative legibility of the small letters. *Amer. J. Psychol.*, 1888, 1, 402–435.
27. SHAW, E. R. School hygiene, Ch. 9. New York, 1908.
28. STARCH, D. Educational psychology, 261–296. New York, 1919.
29. TERRY, P. W. How numerals are read. *Suppl. Educ. Monog.*, 1922, no. 18, 78–93.
30. WELLS, W. Algebra for secondary schools. New York, 1906.
31. WHIPPLE, G. M. Manual of mental and physical tests. Part I: Simple Processes, 305–329. Baltimore, 1914.

THE INFLUENCE OF CONTEXT UPON LEARNING AND RECALL

BY SHUH PAN

Psychological Laboratory, University of Chicago

Introduction

This experiment was designed to test the dependence of learning and recall upon the environmental situation in which the material is presented. Our subjects were required to memorize pairs of words presented in various contextual situations, and they were later tested for their ability to recall this material when the contextual situation was present, absent, or altered in certain specified ways.

A somewhat similar experiment has been performed with the use of animals under the maze problem. After his rats had mastered the problem, Watson (6) rotated the maze in reference to its extraneous environment, and then tested the ability of his animals to run the maze with this change of environmental conditions. The experiment was repeated by Porter (4) and by Hunter (2), using sparrows and pigeons respectively. In all three experiments the animals were confused, the confusion being indicated by an increase in the time and in number of errors. Presumably the animals had learned to respond to some extent to the sensory environment in which the maze was situated, and hence they were disturbed when this relation of the maze to its environment was altered.

This dependence of the 'maze habit' upon the extraneous environment has been systematically studied by Carr (1), who varied the relation of the maze to its environment in various ways while the rats were learning and after they had mastered the maze. His experimental conditions and results are listed in Table I. On the basis of these data Carr concluded that the rat learns to adjust to the entire sensory situation and hence stability of this situation is conducive to learning and to the functioning of the act after it is learned. "Any sensori-

TABLE I

A SUMMARY OF CARR'S EXPERIMENTAL RESULTS

Alteration of Conditions during Recall (after Learning)	Effects
1. Covering maze with canvas top after it was learned	No effect
2. Uncovering maze after learning. Change from poor daylight illumination to normal daylight illumination	No effect
3. Uncovering maze. Change from artificial illumination to daylight illumination.	5 out of 10 rats affected. Average error for those affected increased from 0.20 to 1.07
4. Increase of illumination	7 out of 10 rats affected. Average error increased from 0.51 to 1.35
5. Decrease of illumination	7 out of 10 rats affected. Average error increased from 0.21 to 3.18
6. Change of position of experimenter	All rats affected. Average error increased from 0.11 to 2.50
7. Rotation of uniform environment (canvas top)	No effect
8. Rotation of heterogeneous environment (canvas top with one side open)	5 out of 7 rats affected. Average error, 1.90
9. Change of position of maze in the room	4 out of 6 rats affected. Average error, 2.08
10. Rotation of maze. 5 experiments with various conditions	All showed disturbance. Results in harmony with those secured by Watson
11. Rotation of maze and environment	On average 8 out of 10 rats disturbed. Average error, 1.40
Alteration of Conditions during Learning	**Effects**
1. Maze rotated 90 degrees for each day's test	Average number of trials required for mastering the maze, 30. Group zero error record first obtained on 36th trial Average error, 196. The corresponding values for a group learning the same maze while it is stationary were 18, 22, and 144 respectively. A comparison between two other groups with previous experience upon different types of problems showed similar results.
2. Maze covered by canvas top	Average number of trials, 26; average number errors, 282. The corresponding numbers for the control-group were 18 and 144 respectively

motor act cannot be regarded as an isolated independent function; the act was learned within a wider sensory environment, and it never ceases to be wholly free from those conditions either during or after its development" (p. 291).

Smith and Guthrie (5) mention two experiments in which human subjects were employed. In one of these experiments the subjects learned a list of 10 nonsense syllables in the laboratory, and relearned it 72 hours later in the laboratory; learned a second list in the laboratory and relearned it out of doors; learned a third list out of doors and relearned it in the laboratory, and, finally, learned a fourth list out of doors and relearned it in the same environmental situation. "In eight of the ten subjects there was greater saving in each case where relearning occurred in the same surroundings in which the first learning had taken place. Two subjects showed in one of their four series a greater saving where relearning had occurred under dissimilar conditions. An average of 11.4 per cent. fewer repetitions were required to relearn in similar surroundings" (p. 112). In their second experiment the subjects copied a list of letters on a typewriter the keys of which had been arranged in a random order. "A variable condition consisting of the presence or absence of the odor of oil of peppermint was used." A second practice occurred 24 hours later. "Thirteen subjects were used, and the average saving in the time required to repeat the list was 9.4 per cent. greater when the repetition took place under the condition of the first practice" (p. 113).

Another study with human subjects is reported by Wong and Brown (8), who had two similar groups of subjects solving Yerkes' multiple-choice problems in two rooms; a 'good' room and a 'bad' room. The 'bad' room was an attic room, poorly lighted, dusty, and filled with various objects in disorder. The 'good' room was a well-lighted and well-furnished office. In the given time allotted, the subjects working under the distractive condition solved an average of 5.6 problems per individual, and the average number of trials required per individual was 230. The corresponding records for the other group were 7.8 and 243. The group working under a favorable

environmental condition was thus able to work faster and solve more problems in a given amount of time.

Materials and Method

The materials to be memorized, with one exception, consisted of lists of 20 pairs of logically unrelated words of common usage. Words of one or two syllables were used for the most part. Thirty-two such lists were employed. With a few exceptions, no word was used twice with the same subject.

These lists were presented by means of a memory drum. The time of exposure per pair of words was 3 seconds. Three or four presentations of a list were given in succession, and the amount learned was tested by an immediate written reproduction. The material was thus only partially learned.

The amount retained after 24 hours was tested by a similar written reproduction. In these tests, both immediate and delayed, the stimulus words were presented in an altered temporal order, and the subjects were required to write the response word upon a blank card provided for that purpose. The time for recall was not limited, but it did not exceed 10 seconds a pair.

In a subsidiary experiment, the material consisted of a list of 24 pictures of human faces, each paired with an unfamiliar name. These pictures were taken from a popular magazine, the two sexes being equally represented. Each face and its paired name were presented on a picture post-card. The 24 cards were arranged in a pack, face down. The cards were turned and exposed one by one. The exposure time per card was 3 seconds. The subjects were tested for the amount learned after each presentation of the pack, and the presentations were continued until the material was perfectly memorized.

Retention in this case was tested after 48 hours. To test retention as well as learning, a duplicate pack of cards without the names was employed, and the order in which the cards were presented was altered. The names were orally reproduced. The time for recall was not limited, but it did not exceed 10 seconds a card.

All reproductions for both types of material were scored in terms of the number of verbal associations correctly reproduced. Errors in spelling and pronunciation were ignored in so far as it was evident that the proper word was intended.

Fifty-six subjects served in the experiments, but no subject memorized all of the lists. Thirteen of the subjects were graduate students in the Department of Psychology, and the remainder were Chinese students recruited from various departments.

These materials were presented in a given environmental context. Three kinds of context were employed: words, numbers, and pictures.

The picture on the post-card upon which the face and name were placed constituted the context for this type of material. All of the pictures represented parks, buildings, statues, and other public places in Chicago.

A verbal or numerical context was employed for the pairs of association words. The pair of words to be memorized was always typed in capitals, while the words or numbers constituting their context were always typed in small letters. Usually the context of each pair consisted of two words, one placed above and one below. In a few lists, only one word was employed for the context because of the difficulty of securing suitable words. Only one number was used, and this number corresponded to the position of the pair in the order of presentation. When a number or one word was used, it was written above the pair of associated words. The contextual words bore one of several relations to the pair to be memorized. The kinds of context used are illustrated in Table II. The contextual words may bear no logical relation to either of the members of the pair; they may be logically related to the response word; logically related to the stimulus word; or logically related to both stimulus and response words. Each pair of a list has its own context, and this particular contextual relation always remains constant during the period of learning, unless otherwise specified. The various pairs of associated words composing any list always have the same kind of context.

TABLE II

KINDS OF CONTEXT USED IN THE MAJOR EXPERIMENTS

Kinds of Context	Examples
1. No context	(1) HALF—DEW (2) FOLD—SOUTH
2. Logically unrelated words	lock (1) BEST—CHILL brown steam (2) MAT—EGO court
3. Words logically related to response word	match (1) ZOO—SMOKE pipe knob (2) BIT—DOOR key
4. Words logically related to stimulus word	fire (1) GUN—PRINT hunting wool (2) LAMB—MAGNET mutton
5. Word logically related to stimulus and response words	cane (1) LAME—SUGAR cloth (2) MOTH—FASHION
6. Number	1 (1) RIOT—MELON 3 (2) ACID—HOUSE

The study was carried out in four parts. In Table III are presented the lists used, the number of subjects, the number of presentations given, and the contextual conditions during learning and recall.

In order to ascertain the influence of context upon the rate of learning, we have for comparison (1) lists learned with no context present, (2) lists learned in a given kind of context, and (3) lists in which the particular contextual words were changed after each presentation without any alteration of the kind of context. This latter condition is termed a variable context.

TABLE III

Lists	Condition of Learning	No. of Presentations	Condition of Recall
	PART I. 24 SUBJECTS		
A and *A'*....	Logically unrelated word-context	3	Unaltered context
B and *B'*....	Logically unrelated word-context	"	Removal of context
C and *C'*....	Word-context logically related to response word	"	Unaltered context
D and *D'*....	Word-context logically related to response word	"	Removal of context
E and *E'*....	No context	"	No context
F and *F'*....	No context	"	Introduction of word-context logically related to response word
	PART II. 12 SUBJECTS		
I*a* and I*a'*...	Variable word-context logically related to response word	4	Removal of context
I*b* and I*b'*...	Variable word-context logically related to response word	"	New word-context
II*a* and II*a'*.	Word-context logically related to response word	"	Removal of context
II*b* and II*b'*.	Word-context logically related to response word	"	New word-context
	PART III. 16 SUBJECTS		
1.....	Word-context logically related to response word	3	Context interchanged
2.....	Word-context logically related to response word	"	Context interchanged
3.....	Word-context logically related to stimulus word	"	Unaltered context
4.....	Word-context logically related to stimulus word	"	Removal of context
5.....	Word-context logically related to both stimulus and response words	"	Unaltered context
6.....	Word-context logically related to both stimulus and response words	"	Removal of context
7.....	Number context	"	Unaltered context
8.....	Number context	"	Removal of context

TABLE III—*Continued*

Lists	Condition of Learning	No. of Presentations	Condition of Recall
	PART IV. 12 SUBJECTS		
II1	No context	4	No context
II2	Word-context logically related to response word	"	Unaltered context
II3	Word-context logically related to response word	"	New word-context
III1	No context	3	No context
III2	Word-context logically related to response word	"	Unaltered context
III3	Word-context logically related to response word	"	Context interchanged

To study the effect of context upon the ability to recall, we have for comparison the following conditions: (1) no context present during either learning or recall, (2) the removal of the context that was present during learning, (3) the introduction of a context when none was present during learning, (4) the retention of the context that was present during learning, (5) changing the particular words used without altering the kind of context, and (6) interchanging in recall the contexts for the various pairs.

In the fifth condition listed above, two new context-words are substituted for those employed during learning whose logical relation to the pair is, however, the same as during learning. In the sixth condition, the particular context of each pair during learning is employed as the context of some other pair of the list during recall.

For the pictorial material two groups of subjects were employed. The context remained constant throughout learning. For one group of subjects, one half of the faces were presented in recall in the same context in which they were learned and one half were presented in a new context. Those pairs for which the context was changed for one group of subjects were presented with an unaltered context for the second group, and *vice versa*. These conditions are listed in Table IV.

TABLE IV

	Cards 1, 3, 4, 6, 8, 9, 11, 14, *etc.*	2, 5, 7, 10, 12, 13, 15, *etc.*
Group I	Changed	Not changed
Group II	Not changed	Changed

Before beginning the regular experiments all subjects were required to memorize two lists without a context in order to become familiar with the experiment and to eliminate some of the practice effect. We attempted to equate for further practice effects by requiring each subject of a group to memorize the various lists in a different temporal order. As an illustration, the order used for the first group of subjects in the first part of the experiment is given in Table V.

TABLE V

Subjects	List					
L.	*A*	*B*	*C*	*D*	*E*	*F*
Yu.	*B*	*C*	*D*	*E*	*F*	*A*
Su.	*C*	*D*	*E*	*F*	*A*	*B*
Tz.	*D*	*E*	*F*	*A*	*B*	*C*
Le.	*E*	*F*	*A*	*B*	*C*	*D*
K.	*F*	*A*	*B*	*C*	*D*	*E*

The effect of any differences in the difficulty of the lists was eliminated by using two lists and several groups of subjects for each condition, and by repeating the condition with other lists and subjects in another part of the experiment.

It is obvious that a very definite set of instructions must be employed in order to secure uniform results in such an experiment. Our instructions were typed and presented to the subjects for perusal before they began the experiment. Additional instructions were given them for each change of procedure. An effort was always made to keep the subjects ignorant of the purpose of the experiment. Two typical sets of instructions are appended.

INSTRUCTION USED IN PRACTICE

"This is a memory experiment. The experimenter will expose on the drum a pair of words written in this fashion: MEMORY—BANANA. Twenty such pairs will constitute a series. Your task is to associate each pair of words so as to be able to recall the second word when the first is presented alone. After the series has been presented 8 times, the first word of each pair will be presented in the opening to the right, and you will be expected to write down the corresponding word of the pair on

a blank page furnished for this purpose. In case you are unable to recall the word, leave the space blank. You can write down whatever word is suggested, even though you are not confident of its accuracy. You will also be retested 24 hours later. Please refrain from discussing the experiment or attempting to recall the series during the day."

INSTRUCTION USED IN FIRST PART OF EXPERIMENT

"This experiment will be the same as formerly save in the following respects. Besides the pair of words typed in capitals to be memorized, there may be presented two accessory words, one above and one below the pair to be memorized: *e.g.*,

world
MEMORY—BANANA
fox

These two accessory words will sometimes be present in learning and sometimes absent. They will also be present at times in recall and absent at other times. If they are not present during recall, you are not expected to attempt to write them down. So far as you are concerned, the experiment will proceed just as if these accessory words were not given. You are to memorize the pair of words typed in capitals, and in the test for recall to write down the second word of the pair. You will have three presentations of the series of words."

THE INFLUENCE OF CONTEXT ON LEARNING

The comparative data are given in Tables VI to XI. These tables present the number of subjects in each group, the number of lists learned by each subject, the average number of words learned by each group for each experimental condition, and the average percentage of the words in the list that were learned for each experimental condition. The letters enclosed in parentheses refer to the groups of subjects. The tabular data also include the standard deviation of the mean, and the standard error of the difference of the means.

1. The presence of a word-context that is logically related to the response word exerts a facilitating effect upon learning. The data are given in Table VI. The context increased the number of words learned in all eight of the group comparisons. On the average about two more words per individual were learned under the contextual condition. The difference is not large, but it is consistent.

2. The presence of a word-context logically related to both stimulus and response words also facilitates learning. The data are given in Table VII. The average records of both experimental groups are larger than either of the group records for the control condition. On the average, each indi-

Table VI

Experimental Condition: Word-context Logically Related to the Response Word. Control Condition: No-context. Group Records: Average Number of Words Learned by Different Groups. Per Cent.: Per Cent. of Words in the List Learned. S : L: Number of Subjects in Each Group and Number of Lists Learned by Each Subject. Letters in Parentheses Refer to Various Groups of Subjects. SD: Standard Deviation of the Mean. SED: Standard Error of the Difference of the Means.

Part of Experiment	I		IV		IV	
No. of Presentations	3		4		3	
Condition	Exper.	Control	Exper.	Control	Exper.	Control
S : L	6 : 2	6 : 2	6 : 2	6 : 2	6 : 2	6 : 2
Group Records	(*a*) 15.7 (*b*) 16.9 (*c*) 15.5 (*d*) 16.8	(*a*) 14.1 (*b*) 12.4 (*c*) 13.1 (*d*) 14.3	(*i*) 18.1 (*j*) 16.7	(*i*) 16.6 (*j*) 15.5	(*i*) 16.1 (*j*) 16.2	(*i*) 15.5 (*j*) 14.8
Average	16.2	13.5	17.4	16.1	16.5	15.1
SD	3.4	4.5	2.3	3.6	3.5	4.1
SED	0.8		0.86		1.4	
Per Cent.	81	67	87	80	82	75

vidual learned 2.7 more words under the experimental than under the control condition. The standard error of the difference is 1.2. The difference of the means is 2.2 times larger than the standard error, which means that the chances are 99 in 100 that the true difference is greater than zero.

Table VII

Experimental Condition: Word-context Logically Related to Both Stimulus and Response Words. Control: No-context.
(For Other Explanations, see Table VI)

Part of Experiment	III and IV	
Number of Presentations	3	
Condition	Experimental	Control
S : L	8 : 2	6 : 1
Group Records	(*g*) 17.3	(*i*) 14.8
	(*h*) 18.2	(*j*) 15.5
Average	17.8	15.1
SD	2.2	4.1
SED	1.2	
Per Cent.	89	75

3. A word-context logically related to the stimulus word exerts a detrimental effect upon learning. The data may be found in Table VIII. Both groups working under the exper-

Table VIII

Experimental Condition: Word-context Logically Related to the Stimulus Word. Control: No-context. (For Other Explanations, see Table VI)

Part of Experiment	III	IV
No. of Presentations	3	
Condition	Experimental	Control
S : L	8 : 2	6 : 1
Group Records	(*g*) 11.4	(*i*) 14.8
	(*h*) 13.7	(*j*) 15.5
Average	12.5	15.1
SD	5.3	4.1
SED	1.5	
Per Cent	62	75

imental conditions made poorer records than did either of those for the control condition. When the context was present, an average of 2.6 (13 per cent.) fewer words was learned than when no context was present. The average difference is 1.7 times larger than the standard error of the difference (1.5). The chances are 96 in 100 that the difference is statistically significant.

Table IX

Experimental Condition: Word-context Logically Unrelated to Both Stimulus and Response Words. Control: No-context. (For Other Explanations, see Table VI)

Part of Experiment	I	
No. of Presentations	3	
Condition	Experimental	Control
S : L	6 : 2	6 : 2
Group Records	(*a*) 14.2	(*a*) 14.1
	(*b*) 9.8	(*b*) 12.1
	(*c*) 9.4	(*c*) 13.1
	(*d*) 13.8	(*d*) 14.3
Average	11.8	13.5
SD	4.8	4.5
SED	0.94	
Per Cent	59	67

4. The presence of a word-context that is logically unrelated to both stimulus and response words also seems to be detrimental to learning. In Table IX are given the results. On the average, 1.7 (8 per cent.) fewer words were learned with the context than without. The detrimental effect of the context is shown in three out of four comparisons. With group 'a' the experimental condition seems to have no effect. This fact, however, can be explained by referring to the individual records, which show that two of the subjects gave each an unusually low score in learning one of the two lists under the control condition. Their scores are 9 and 6. With the other list, their scores are 17 and 13, respectively. Those two unusually low scores are probably responsible for the seeming absence of the effect of the experimental condition in the group-results.

5. The data presented in Table X show the influence of a

TABLE X

EXPERIMENTAL CONDITION: NUMBER-CONTEXT. CONTROL: NO-CONTEXT.
(FOR OTHER EXPLANATIONS, SEE TABLE VI)

Part of Experiment	III	IV
No. of Presentations	3	
Condition	Experimental	Control
S : L	8 : 2	6 : 1
Group Records	(*g*) 12.5	(*i*) 14.8
	(*h*) 15.2	(*j*) 15.5
Average	13.8	15.1
SD	5.5	4.1
SED	1.5	
Per Cent.	69	75

number-context upon learning. The presence of the context decreases the average number of words learned. The decrease is 1.3 words or 6 per cent. The ratio of the average difference to its standard error (1.5) is 0.87, which indicates a chance of 4 to 1 that the difference is significant.

6. In Table XI are given the results showing the influence of a variable word-context which is logically related to the response word. Apparently such a variable word-context is not so favorable to learning as is a constant context of the same kind. Both groups working under the experimental

condition learned fewer words than did either of those working under the constant contextual condition. On the average, the difference is 1.7 words. This difference is 3.4 times larger than its standard error (0.5), and it is statistically significant.

TABLE XI

EXPERIMENTAL CONDITION: VARIABLE WORD-CONTEXT LOGICALLY RELATED TO THE RESPONSE WORD. CONTROL 1: CONSTANT WORD-CONTEXT LOGICALLY RELATED TO THE RESPONSE WORD. CONTROL 2: NO-CONTEXT. (FOR OTHER EXPLANATIONS, SEE TABLE VI)

Part of Experiment	III		IV	
No. of Presentations	4			
Condition	Exper.	Control 1		Control 2
S : L	8 : 2	6 : 2		6 : 1
Group Records	(*e*) 16.9	(*i*) 18.9		(*i*) 15.5
	(*f*) 17.3	(*j*) 18.6		(*j*) 16.6
Average	17.1	18.8		16.1
SD	2.4	1.4		3.6
SED	0.5		1.1	
Per Cent	85	94		80

The variableness of the context, however, has not obliterated entirely the facilitating effect of the context which would be shown if it were constant. Both experimental groups made a better record than did either of those for control 2, in which no context was present. On the average, the difference is 1.0 in favor of the experimental groups. Its standard error is 1.1, which indicates that the chances are 82 in 100 that the difference is reliable.

7. On the whole, three kinds of context are found to be beneficial to learning: (1) a word-context logically related to the response word, (2) a word-context logically related to both stimulus and response words, and (3) a variable word-context logically related to the response word. The last contextual condition is least effective in facilitating learning. The effectiveness of the first two kinds of context is nearly equal.

The effect upon learning of the other three kinds of context studied is detrimental. Most detrimental is a word-context logically related to the stimulus word. The number-context

and a word-context logically unrelated to both stimulus and response words are nearly on a par in this respect.

8. In Fig. 1 is shown a learning curve from the experiment dealing with the association of pictures of faces with names. The data have nothing to do with our present prob-

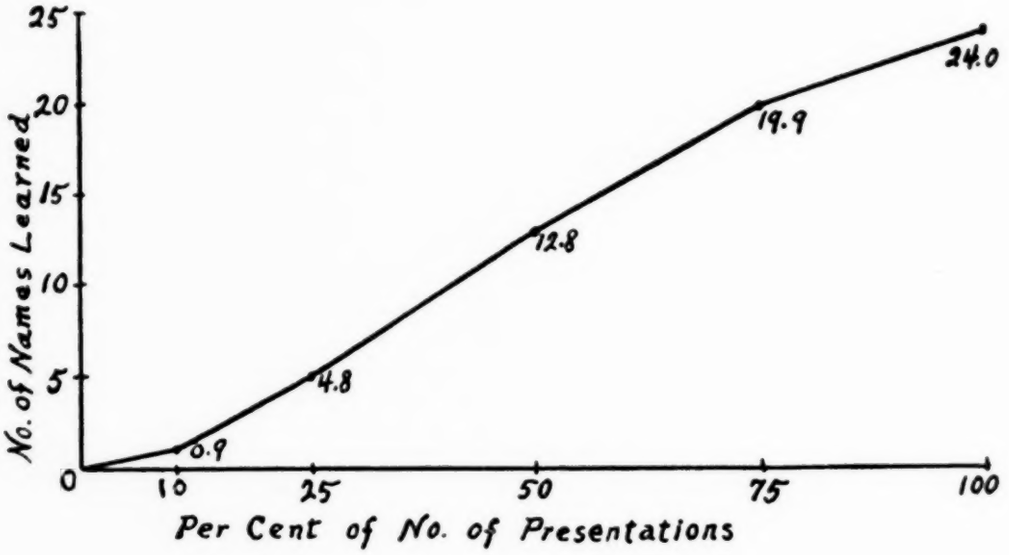

FIG. 1. A composite learning curve of 12 subjects on the association of names with pictures of human faces.

lem. They are recorded merely because of the unusual character of the results. The curve is a composite curve for 12 subjects, plotted according to the method first used by Kjerstad (4). It exhibits a slight positive acceleration at the beginning and a slight negative acceleration at the end. Most of the individual curves show the same characteristics.

THE INFLUENCE OF CONTEXT ON RECALL

The comparative data showing the effect of the various contextual conditions on recall are given in Tables XII–XVI, and in Fig. 2. The arrangement of the tables is similar to that given in the previous section.

1. The results showing the effect of the removal of a context in recall are given in Table XII. In three out of five cases the removal of the context is detrimental, or, conversely, one can recall most readily under those contextual conditions in which the material was learned. The detri-

mental effect is greatest when the verbal context is logically related to the response word. It is also considerable when the verbal context is logically related to both stimulus and response words. With the number-context, the detrimental

TABLE XII

EFFECT OF REMOVAL OF A CONTEXT. AVE. NO. WORDS AND PERCENTAGE RECALLED UNDER SPECIFIED CONDITIONS. PART OF EXPER.: THE PART OF THE EXPERIMENT IN WHICH THE RESULTS WERE SECURED. S : L: NUMBER OF SUBJECTS IN EACH GROUP AND THE NUMBER OF LISTS RECALLED BY EACH SUBJECT. GROUP RECORDS: AVE. NO. WORDS RECALLED BY VARIOUS GROUPS. LETTERS IN PARENTHESES REFER TO GROUPS. SD: STANDARD DEVIATION. SED: STANDARD ERROR OF DIFFERENCE.

Part of Exper.	I		I	
Learning Condition	Word-context related to response word		Word-context unrelated to stimulus and response words	
Recall Condition	Unaltered context	Context removed	Unaltered context	Context removed
S : L	6 : 1	6 : 1	6 : 1	6 : 1
Group Records	(*a*) 11.3 (*b*) 12.7 (*c*) 9.5 (*d*) 14.5	(*a*) 3.5 (*b*) 2.7 (*c*) 1.2 (*d*) 3.6	(*a*) 4.5 (*b*) 3.3 (*c*) 2.3 (*d*) 6.0	(*a*) 6.5 (*b*) 2.5 (*c*) 3.3 (*d*) 7.3
Average	12.0	2.7	4.0	4.9
SD	4.6	3.1	5.0	4.0
SED	1.1		1.3	
Per Cent.	75	16	33	39

Part of Exper.	III		III		III	
Learning Condition	Word-context related to stimulus word		Word-context related to stimulus and response words		Number-context	
Recall Condition	Unalt'd context	Context removed	Unalt'd context	Context removed	Unalt'd context	Context removed
S : L	8 : 1	8 : 1	8 : 1	8 : 1	8 : 1	8 : 1
Group Records	(*g*) 4.4 (*h*) 4.6	(*g*) 4.5 (*h*) 4.3	(*g*) 12.9 (*h*) 12.2	(*g*) 8.1 (*h*) 7.1	(*g*) 7.8 (*h*) 6.9	(*g*) 6.3 (*h*) 5.2
Average	4.5	4.4	12.5	7.6	7.3	5.7
SD	3.9	4.6	4.1	5.3	5.0	4.4
SED	1.5		1.7		1.5	
Per Cent.	37	33	70	43	48	43

effect is slight but appreciable. The probabilities are 5.6 to 1 that the difference between the means is reliable.

The removal of the context logically related to the stimulus word apparently does not exert any effect upon the ability to recall.

The removal of the unrelated word-context apparently exerts a slight beneficial effect upon recall. The differences between the means are small and not consistent. The chances are but 3 to 1 that the difference is significant. In other words, the presence of the contextual conditions under which the material was learned exerts a very slight distractive effect upon the ability to recall.

2. In recall the introduction of a word-context logically related to the response word proves to be beneficial. The data are presented in Table XIII. On the average, the introduction of such a context increased the amount recalled from 42 per cent. to 62 per cent. The results are consistent for all groups and the differences are large enough to be statistically significant.

TABLE XIII

THE EFFECT ON RECALL OF INTRODUCING A CONTEXT
(FOR EXPLANATIONS, SEE TABLE XII)

Part of Exper.	I	
Learning Condition	No context	
Recall Condition	No context	Introduction of word-context logically related to response word
S : L	6 : 1	6 : 1
Group Records	(*a*) 6.5	(*a*) 9.1
	(*b*) 4.1	(*b*) 8.5
	(*c*) 3.5	(*c*) 8.7
	(*d*) 7.5	(*d*) 10.3
Average	5.4	9.2
SD	4.1	4.5
SED	1.2	
Per Cent.	42	62

3. The change from the original context to a new context of the same kind is distinctly unfavorable to recall.

In the case of a word-context logically related to the re-

sponse word, the detrimental effect of such a change is quite pronounced, as is shown by Table XIV. Such a procedure,

Table XIV

The Effect of a New Context on Recall.
(For Explanations, see Table XII)

Part of Exper.	IV	
Learning Condition	Word-context logically related to response word	
Recall Condition	Unaltered context	New context
S : L	6 : 1	6 : 1
Group Records	(*i*) 15.1	(*i*) 6.8
	(*j*) 15.1	(*j*) 6.1
Average	15.1	6.5
SD	2.7	3.1
SED	1.2	
Per Cent.	87	37

however, introduces two changes: namely, the removal of the original context and the introduction of the new context. It has been shown in Table XII that with the removal of a word-context logically related to the response word only an average of 16 per cent. was recalled. In the present case, in which we have a new context in place of the original, the amount recalled is 37 per cent., or an increase of about 20 per cent. The new context, therefore, exerts a beneficial effect as compared with the removal of the original context. In the preceding section we have noticed that the introduction of such a word-context when none was present during learning increased the amount recalled from 42 per cent. to 62 per cent.; also an increase of about 20 per cent. Thus, the presence of such a new context in recall seems to exert a definite amount of facilitating influence, whether by itself or in combination with some other sort of change.

In the case of the association of pictures of faces with names, the detrimental effect of a new context is less pronounced, but consistent. The results are presented in Table XV. Of the 12 subjects, 9 recalled fewer names under the new contextual condition than when the context was unaltered. The average decrease is from 57 per cent. to 40 per cent. with Group I and from 72 per cent. to 50 per cent. with Group II.

Table XV

The Effect of a New Context on Recall in the Case of the Association of Pictures of Faces with Names

							Total	Ave.	Per Cent.
		Group I							
Subject	*A*	*B*	*C*	*D*	*E*	*F*			
No. of names recalled with new context	3	0	3	4	5	9	24	4	40
No. of names recalled with unaltered context	6	2	8	5	6	7	34	5.7	57
SED								1.3	
		Group II							
Subject	*O*	*P*	*Q*	*R*	*S*	*T*			
No. of names recalled with new context	3	8	7	3	4	5	30	5	50
No. of names recalled with unaltered context	8	6	5	7	7	10	43	7.2	72
SED								0.97	

4. The comparative data showing the influence of the interchange of word-contexts which are logically related to the response words are presented in Table XVI. Such inter-

Table XVI

The Effect on Recall of Interchange of Context
(For explanations, see Table XII)

Part of Exper.		IV	
Learning Condition		Word-context logically related to response word	
Recall Condition		Unaltered context	Context interchanged
S : L		6 : 1	6 : 1
Group Records	(*i*)	9.8	(*i*) 1.5
	(*j*)	10.5	(*j*) 1.8
Average		10.2	1.6
SD		3.1	3.9
SED		1.4	
Per Cent.		66	9

change of context in recall is very detrimental. With the context unaltered, an average of 66 per cent. was recalled; while with contexts interchanged only 9 per cent. was recalled. The detrimental effect is even greater than that due to the removal of such a context as is shown in Table XII. In the individual records we found that many words were recalled which were correct so far as the contexts were concerned, but

which were wrong in respect to the stimulus words with which they were associated.

5. The disintegrative effect of a change of context in recall varies with the degree of learning.

With the picture material, the subject was tested for the amount learned after each presentation of the series. Some pairs were learned in the early trials, some in the later trials, and others in the final trials. Those pairs that were learned in the early trials naturally received a number of additional presentations and were overlearned as compared with those that were learned in the last trials. In other words, the degree to which the various pairs of associates were overlearned varied directly with the order in which they were learned. We divided each subject's learning into three approximately

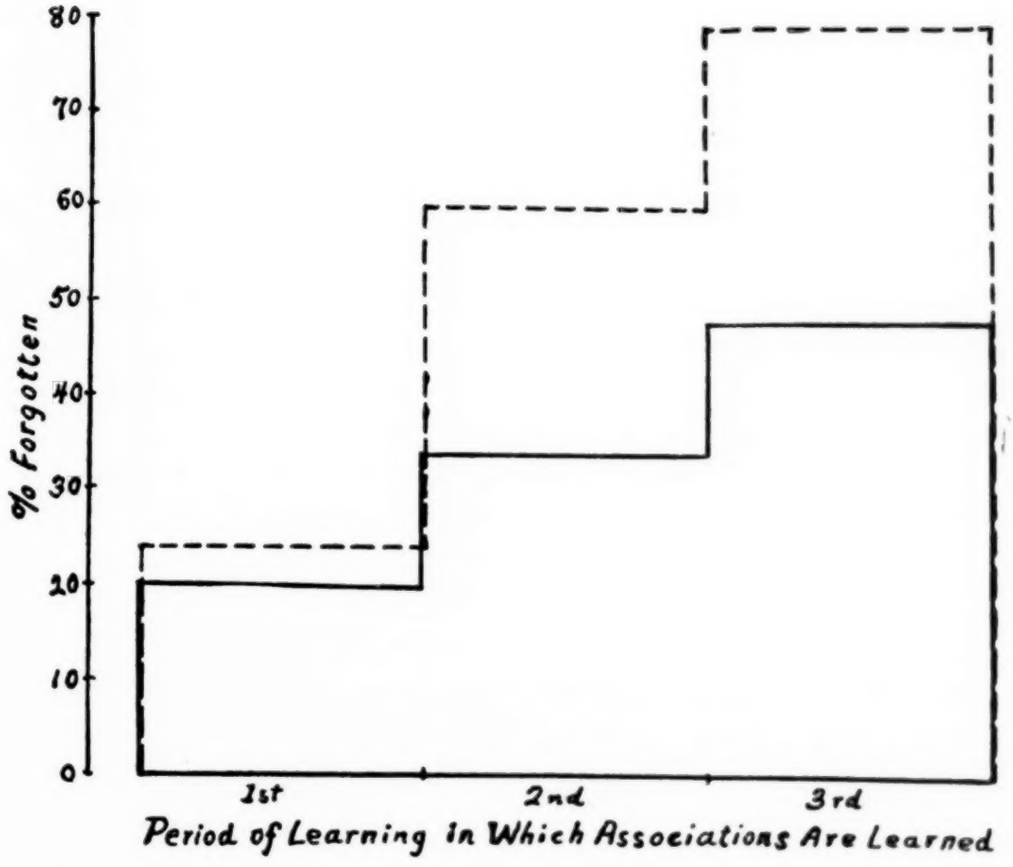

FIG. 2. The relation between degree of learning and susceptibility to effect of change of context for picture-name associations. Results of 12 subjects. Dotted line represents amount forgotten with context changed. Solid line represents amount forgotten with context unchanged.

equal periods, and grouped the pairs of associates according to the period in which they were learned; or, in other words, in respect to the degree to which they were memorized. The first group to be memorized represents the highest degree of learning, and the last group memorized represents the least degree of learning. In order to ascertain the relation between the disintegrative effect of a change of contextual conditions to the degree of learning, we computed the amount forgotten under the changed and unchanged contextual conditions for each of the three groups of associates.

The comparative data are graphically represented in Fig. 2. With the unaltered context, the amounts forgotten were 20 per cent., 34 per cent. and 48 per cent. for the first, second and third groups respectively. With the altered context, the corresponding values were 24 per cent., 60 per cent. and 79 per cent. Under both conditions, the amount forgotten varies inversely with the degree of learning. For all three groups, or for all degrees of learning, more was forgotten when the context was altered, but the difference between the two contextual conditions increases as the degree of learning decreases. In other words, a change of contextual conditions exerted the most disintegrative effect upon the material which was the least learned.

Conclusions

Both learning and recall were influenced by the contextual conditions. The nature and amount of the effect varied with the nature of the conditions.

A word-context logically related to the response word exerted a beneficial effect upon learning, and its removal or interchange during recall was highly detrimental. When the context was varied during learning, the effect was lessened. The introduction of such a context in recall when none was present in learning also exerted a facilitating effect. The substitution in recall of a new context of the same kind exerted about the same effect as the introduction of such a context when none was present during learning.

A word-context logically related to both the stimulus and

the response words facilitated learning to about the same degree as did one that was logically related only to the response word. Its removal in recall, however, exerted a less detrimental effect than did the latter.

A word-context logically related to the stimulus word exerted a detrimental effect upon learning. Its removal in recall did not exhibit any appreciable influence.

A word-context logically unrelated to both the stimulus and the response words was also detrimental to learning. Its removal in recall was apparently beneficial.

The presence of a number-context during learning was slightly detrimental, and its removal in recall exerted a similar effect.

The substitution of a new context of the same kind (picture post-cards) also exerted a detrimental effect upon the recall of the face and name associations. The degree of the effect varied inversely with the degree of learning.

The explanation of these results is obvious in most cases.

A logical relation between two words presupposes some sort of *a priori* associative connection between them, so that the presence of one naturally tends to arouse the other. The introduction in recall of a word-context that is logically related to the response word, although it was not present in learning, will facilitate the arousal of the proper associate because both the stimulus-word and the context-word are associated with it. Likewise, the presence of such a contextual word during learning will exert a facilitating effect upon the learning process, and naturally the retention of this word in recall will be distinctly favorable. Even the variation of this type of context during learning and recall will exert some beneficial effect. The interchange of such a context in recall will exert a detrimental effect, for in any given situation the stimulus-word will tend to arouse one associate and the context-word another.

When the context-word is unrelated to either the stimulus or response words, the subjects may ignore it; their attention may be distracted from the task in hand, or they may attempt to establish a double association. In the two latter cases, the

context will presumably exert a detrimental effect. As a matter of fact, this context did interfere with learning to some extent, and its presence in recall may have been slightly detrimental.

Practically the same results were obtained when the context-word was logically related only to the stimulus-word. Apparently, the presence or absence of an associative connection between the stimulus- and the context-words does not materially influence the subjects' reactions to the situation.

From the preceding results, we should expect a context that is related to both the stimulus-word and the response word to exert an effect like that of a context that is related only to the response word, for a connection between the context and the stimulus-word does not seem to be effective. The results justify this assumption. This type of context is distinctly favorable to learning, and its presence in recall is also quite effective. In fact, the ability to recall seems to be more dependent upon this type of context than upon one that is related only to the response word.

The number-context was detrimental to learning but beneficial in recall. We must, as it appears, assume that the attempt to utilize these clues in the initial presentations was distractive; while the subjects were able to employ them with some degree of success in the final presentation, so that their presence in recall was beneficial.

Our results thus indicate that the recall of any material is favored by the presence of an environmental factor which has some associative connection with that material. In the absence of such an association, the environmental situation is likely to be unfavorable to recall. When the material is learned in a contextual situation with which it has no associative connection, such an association is likely to be established with a sufficient number of presentations. Such an association was established in the case of the face-name material, and the strength of the connection increased with the number of the presentations. In the case of the unrelated verbal context, three presentations were apparently insufficient to form an effective connection.

REFERENCES

1. CARR, H. Maze studies with the white rat. I. Normal animals, *J. Animal Behav.*, 1917, 7, 259–75.
2. HUNTER, W. S. Some labyrinth habits of the domestic pigeon, *J. Animal Behav.*, 1911, 1, 278–304.
3. KJERSTAD, C. L. The form of the learning curves for memory, *Psychol. Monog.*, 1919, **26** (no. 116).
4. PORTER, J. P. Further study of the English sparrow and other birds, *Amer. J. Psychol.*, 1906, 17, 248–71.
5. SMITH, S., & GUTHRIE, E. R. *General psychology in terms of behavior*, New York, 1921.
6. Watson, J. B. Kinaesthetic and organic sensations; their role in the reactions of the white rat to the maze, *Psychol. Monog.*, 1907, 8 (no. 33).
7. WATSON, J. B. Behavior of noddy and sooty terns, *Carnegie Instit. Wash. Publ.*, 1909, no. 103.
8. WONG, H., & BROWN, W. Effects of surroundings upon mental work as measured by Yerkes multiple choice method, *J. Comp. Psychol.*, 1923, 3, 319–31.

VARIABILITY OF JUDGMENTS ON MUSICAL INTERVALS

BY HELEN MORAN AND CARROLL C. PRATT

Harvard Psychological Laboratory

A musical interval, or any separation of two impressions on the pitch-continuum, is generally defined or recorded in terms of the ratio of the vibration-frequencies of the two generating physical bodies. A fifth, *e.g.*, is that interval whose components stand in the ratio 2/3; a minor sixth has the ratio 5/8, *etc.* But these definitions are wholly in terms of physical properties. Nothing has been said about the interval considered as a datum of psychological experience. There is not a great deal, in point of fact, which can be said as yet concerning the phenomenological aspects of auditory intervals. Stumpf singled out certain aspects of musical intervals for exhaustive study and detailed discussion. Among other descriptive matters he observed differences among them with respect to the degree to which they approximate a single auditory impression, and he found it possible to work out uniformities in the degrees of *fusion* among intervals. But Stumpf was not always clear as to just what he meant by fusion.[1] And later investigations [2] have shown that observers may pass comparative judgments on intervals under at least half a dozen totally different criteria of judgment, which would seem to indicate that in spite of whatever uniformities may be secured in the order of intervals when judged by degrees of fusion, there are other characteristics, such as smoothness, simplicity and affective tone, which must be considered as entering into, or at least as related in some way to, the

[1] For ambiguities in Stumpf's definitions of fusion see *Tonpsychologie*, 1890, vol. 1, 127, 128; *Konsonanz und Dissonanz*, 1898, 35, 44; and *cf.* M. Bentley, *Amer. J. Psychol.*, 1903, **14**, 65–70.

[2] W. Kemp, Methodisches und Experimentelles zur Lehre der Tonverschmelzung, *Arch. f. d. ges. Psychol.*, 1913, **29**, 139–181; C. C. Pratt, Some Qualitative Aspects of Bitonal Complexes, *Amer. J. Psychol.*, 1921, **32**, 490–515.

experience of fusion. It is not surprising, then, that fusion is difficult to define with precision.

A recent attempt has been made [3] to discover whether each musical interval may not be characterized by certain phenomenological aspects peculiar to it alone. It was found that most of the observers were able, after preliminary practice, to find terms, drawn largely from taste and touch, which served adequately to characterize and to differentiate the various intervals. It is probable that such phenomenological characteristics form the original basis for the quick recognition of the various musical intervals possessed by many musicians. It is not uncommon to find that a musically trained person can distinguish, say, a fifth from a minor third about as readily as people tell red from green, although in most cases the person is quite unable to say exactly how he makes the distinction.

Since it appears, then, that each interval has a phenomenological aspect peculiar to itself, and since many musical persons have absolute recognition of intervals, the question arises as to the range over which a given characterization applies, or, in other words, to the accuracy with which recognition of a given interval takes place. How far may an interval formed by two simultaneous tones depart from the ratio, say, 2/3 and still be characterized as a fifth? For the physicist a fifth would be that interval which has exactly the ratio 2/3; but "fifthness," regarded as a psychological datum, may be correlated with a physical range which extends some distance both above and below the 2/3 ratio. It was with the purpose of finding the range over which two tones of different pitch may vary without altering the psychological characteristics of the interval that the present study was undertaken.

There are reports in the literature of various determinations of the j.n.d. of musical intervals.[4] These values are chiefly derived either from few observations from very different methods, or from tones presented successively. They are therefore hardly comparable with our own. Stumpf has discussed the problem at length,[5]

[3] E. M. Edmonds & M. E. Smith, The phenomenological description of musical intervals, *Amer. J. Psychol.*, 1923, **34**, 287–291.

[4] References to this literature and a summary of the principal results are given by Titchener in his *Experimental Psychology*, 1905, Vol. II, pt. ii, 244–248.

[5] C. Stumpf & M. Meyer, 'Maassbestimmungen über die Reinheit consonanter

and (with Meyer) has presented the results of observations on the octave, fifth, major third, and minor third. In the few trials with simultaneous tones, j.n.d.'s were secured which (if we may here compare results) are of about the same magnitude as the crude errors of our values.[6] Most of the experiments, however, were made with successive tones, and the discrimination of intervals generally turns out to be finer for successive than for simultaneous presentations.[7]

The present experiments were done during the academic year 1924–25. There were, fortunately, three Os in the laboratory at that time who were trained in psychological observation and who also were able to recognize more or less immediately any given musical interval: Helen K. Mull, Marjory B. Pratt, and C. C. Pratt. Two Stern variators (300–600, 400–800 d.v.) were used as stimuli throughout the experiment. They were placed side by side on a table before E. An extension rod, running through the wall of the room, was attached at one end to the pinion gear of one of the variators. At the other end was a handle which could be grasped and easily turned by O, so that the latter, sitting in a room next to the experimenting room, had only to turn the handle to the right to raise the pitch and to the left to lower it. The procedure adopted was essentially the Fechnerian method of average error. One of the variators was always controlled by O after E had set it roughly in a given region. The other variator, after it had been set exactly by E for each trial, was left unchanged. The task of O was to alter the pitch of the upper variator which he himself controlled until he was satisfied that the two tones corresponded to the required interval. His instructions were made to read as follows:

"In this experiment you will be asked to vary one of two tones until the two of them together are as close to a given musical interval as you can get them. The pitch is raised by turning the handle to the right, and lowered by turning it to the left. Before each trial E will announce to you the name of the interval to which you are to set the tones. Following the 'ready' signal you will hear the two tones. You will then vary the upper one until the two tones sound like the interval asked for. Tell E as soon as you are satisfied that the two tones approximate as closely as possible the required interval."

During the second part of the experiment the *lower* tone was

Intervalle,' *Beiträge zur Akustik und Musikwissenschaft*, 1898, 2, 84–167.

[6] *Ibid.*, 123, 130.

[7] *Ibid.*, 129 f.

placed under the control of O and the instructions changed accordingly. In both parts of the experiment the lower tones ranged between 345 and 495 d.v., and the higher tones between 520 and 655. The intervals were presented to O in haphazard order. Forty judgments were made on each of eleven intervals (all but the octave): twenty judgments in the first part when the upper tone was varied, and twenty in the second part when the lower tone was varied. In order to overcome the error of direction E set the variable tone half the time above the correct position and half the time below. These changes in direction were arranged in haphazard fashion so that O could not know in advance whether he would be called upon to lower or to raise the variable tone.

The Os were asked occasionally to give introspective characterizations of their experience. These reports yielded little of importance. The recognition of the intervals was, in general, quick and fairly mechanized. The reports did show, however, that the procedure for setting the intervals was about the same for all three Os. It was necessary, first, to take the interval as a whole; then, having decided whether it was too large or too small, to abstract from the whole to the variable tone while the latter was undergoing its alteration; to return again to the interval as a whole for another judgment; to abstract again to the single tone if the interval was still untrue; and so on, until O was satisfied that the two tones were as close as he could get them to the required interval.

Examination of the results reveals at a glance the fact that the range of physical values over which a given interval may vary psychologically is very small. In Table I are presented the mean variations of the average settings of the Os for each of the eleven intervals, as well as the differences between these settings and the natural and tempered intervals.

In order to facilitate comparison of m.v.'s and differences from various parts of the tonal scale, certain of the data of Table I have been changed from absolute to relative values by converting the ratios of vibration-frequencies into cyclic cents.[8] One cent equals one-hundredth of an equal semitone,

[8] For a discussion of the cent and its calculation, see appendix of A. J. Ellis' translation (2nd ed., 1885) of Helmholtz, *Die Lehre von den Tonempfindungen, usw.*, p. 437, art. 24, xiii, and pp. 446–457.

TABLE I

JUDGMENTS ON MUSICAL INTERVALS BY METHOD OF AVERAGE ERROR. M.V. = MEAN VARIATION OF AVERAGE SETTING BY O; NAT. = DIFFERENCE BETWEEN AVERAGE SETTING AND NATURAL INTERVAL; TEM. = DIFFERENCE BETWEEN AVERAGE SETTING AND TEMPERED INTERVAL

	Series I (upper tone variable) Range of upper tones: 524–652 d.v. Range of lower tones: 345–495 d.v.									Series II (lower tone variable) Range of lower tones: 347–492 d.v. Range of upper tones: 525–650 d.v.								
Observer	HM			MP			CP			HM			MP			CP		
Interval	M.v.	Nat.	Tem.	M.v.	Nat.	Tem.	M.v.	Nat.	Tem.	M.v.	Nat.	Tem.	M.v.	Nat.	Tem.	M.v.	Nat.	Tem.
Major seventh	7.04	12.02	7.62	10.85	12.99	8.61	13.40	6.11	1.71	2.97	0.45	1.88	4.79	0.32	2.66	2.12	1.30	1.03
Minor seventh	7.91	4.81	3.39	7.45	6.32	4.90	6.07	3.86	5.28	4.56	2.84	2.04	4.88	1.81	1.01	2.94	1.45	0.65
Major sixth	4.44	9.44	3.93	8.35	6.79	1.28	4.59	4.34	1.16	4.50	4.88	0.59	6.14	7.46	4.17	2.80	4.56	1.27
Minor sixth	8.01	0.96	5.62	6.31	1.21	5.87	7.96	3.04	1.62	2.38	0.70	2.22	4.52	1.03	3.95	3.15	4.28	1.36
Fifth	5.11	3.04	3.68	6.02	2.72	3.36	5.60	0.27	0.91	2.76	2.12	1.70	4.42	1.54	1.96	1.30	0.18	0.60
Tritone	7.99	11.15	8.06	8.71	10.42	7.33	7.89	7.06	3.97	3.56	1.70	0.47	6.67	6.84	4.66	2.73	1.51	0.66
Fourth	5.03	0.62	1.25	5.84	6.37	5.64	3.18	0.13	0.76	1.88	1.47	1.95	5.18	0.10	0.58	1.74	0.02	0.46
Major third	6.70	3.30	0.95	7.81	3.33	0.92	1.90	4.38	0.12	3.93	3.15	0.01	3.00	4.41	1.25	2.67	2.58	0.57
Minor third	3.55	4.67	0.18	8.89	5.86	10.72	4.52	3.25	2.60	2.57	3.58	0.45	5.88	5.11	1.07	3.47	5.24	1.20
Major second	4.24	5.91	4.71	10.40	4.16	2.96	6.56	5.74	5.74	3.59	0.25	0.81	4.28	4.32	5.38	3.63	2.29	1.23
Minor second	6.58	0.43	3.16	8.40	2.27	5.86	6.79	3.24	0.34	3.99	3.50	0.14	4.93	5.14	8.50	4.39	3.35	0.01
Averages	6.05	5.12	3.87	8.09	5.67	5.22	6.22	3.76	2.20	3.33	2.24	1.11	4.97	3.46	3.19	2.81	2.43	0.82

or one twelve-hundredth of an octave. These figures are given in Table II.

TABLE II

MEAN VARIATIONS OF TABLE I (CONSIDERED AS MEASURES OF SENSITIVITY IN JUDGMENTS ON MUSICAL INTERVALS) CONVERTED INTO CYCLIC CENTS (SEE TEXT) FOR THE SAKE OF COMPARISON OF VIBRATION-FREQUENCIES FROM VARIOUS PARTS OF THE TONAL SCALE

	Series I			Series II			Averages
	HM	MP	CP	HM	MP	CP	
Major seventh	19	28	35	15	24	11	22
Minor seventh	21	20	16	22	24	14	19.5
Major sixth	13	23	13	21	29	13	17
Minor sixth	23	18	23	11	21	15	18.5
Fifth	15	18	17	13	20	6	14.8
Tritone	25	26	24	16	30	12	22.1
Fourth	16	18	10	8	22	7	13.5
Major third	21	25	6	16	12	11	15.1
Minor third	11	28	15	10	22	13	16.5
Major second	14	33	21	12	16	13	18.1
Minor second	21	27	22	14	17	15	19.3
Averages	18	24	18.3	14.3	21.5	11.8	

These data would seem to indicate that for musical individuals an interval is confined, psychologically, to a narrow range. Of the sixty-six m.v.'s of Table II only eight are greater than 25 cents (one-quarter of an equal semitone). The average m.v. for the three Os for all the intervals is 18 cents, or less than one-fifth of a semitone.

The crude errors, as measured by the difference between the observed settings and the values of the natural and tempered intervals, give indication of still finer discrimination on the part of the Os. In Table III the average values of these differences have been converted into cents. It will be noted that the deviation of the Os' settings from the natural intervals is, on the average, less than one-seventh of a semitone, and from the tempered intervals less than one-tenth of a semitone. One would be inclined to suppose that the difference here in favor of the tempered intervals is due to the fact that musical individuals are generally more familiar with tempered than with natural intervals. It is the mean variation of the average

setting, however, rather than the crude error, which should be taken as a measure of sensitivity. One may say, then, according to this measure, that the average limen for the discrimination of intervals, under the conditions of the present experiment, is about one-fifth of an equal semitone.

Table III

Averages, Expressed in Terms of Cents, of All the Differences between the Average Settings by O for Each Interval and the Physical Values of the Corresponding Natural and Tempered Intervals

	Natural	Tempered
HM	12.5	8.5
MP	16	15
CP	10.5	5.5

The size of this limen has a bearing on the problem of the status of those intervals which lie between the musical intervals. Meyer has been quoted in a recent article [9] as stating that the tempered scale functions exactly as does the natural scale, because the tempered scale in no interval deviates more than 16 cents from the natural scale. Sixteen cents, however, is pretty close to our average limen of 18 cents. And several of the individual limens are appreciably less than the average value. The average m.v.'s for HM and CP are in every case 18 cents or less. (In the second series they are 14.3 and 11.8 respectively.) Of the sixty-six m.v.'s in Table II for the separate intervals, over 50 per cent. are 18 cents or less. Over 25 per cent. are 13 cents or less. For CP the limens for the minor third, major sixth, and major third are 14, 13, and 8.5 cents respectively. For these same intervals the differences between natural and tempered intonation are 16, 16, and 14 cents respectively. HM's limen for the minor third is 10.5 cents. Meyer equates the tempered and natural scales, as far as psychological function is involved, on the assumption that when an interval deviates by not more than a *third* of a semitone it is always assimilated immediately to the nearest

[9] Max Meyer, 'La musica bicromatica, el teclado para ejecucion y la pauta en que escribirla,' reported by Jesus C. Romero, *La Antorcha*, Mexico, 1925, 1, no. 16, 7 f., 34 ff. *Cf.* P. R. Farnsworth, this Journal, 1926, **9**, 255, who accepts Meyer's *dictum*.

familiar interval. But according to our results assimilation takes place only when an interval deviates from some 'norm' by not more than a *fifth* of a semitone. Deviation beyond a fifth of a semitone would not necessarily place an interval in a position which would demand assimilation either to the interval above or below. Such a position is not equivocal unless by equivocality one means merely inability to give a name to the interval. Our Os could only name such an interval in some such fashion as this: "It is not a fifth, nor a minor sixth. It's somewhere in between!" Each interval of the tempered scale is separated from its adjacent intervals by 100 cents. According to our results each interval of a 'psychophysical scale' would have an 'interval of uncertainty' of about 40 cents.

In the last column of Table II are given the averages of the m.v.'s of the three Os for each of the eleven intervals. When these values are ranked from smallest to largest, the order of the intervals is as follows: fourth, fifth, major third, minor third, major sixth, major second, minor sixth, minor second, minor seventh, major seventh, tritone. This order conforms roughly to the order of the degrees of fusion. The difference between any two adjacent intervals is small, but since the whole series of differences fits the schema of fusion, one is inclined to attach some significance to the order.[10]

These results are derived from judgments made by musical Os. It is necessary to have such Os where one wishes to make use of the method of average error. Comparison of intervals, for the purpose of determining some sort of j.n.d., presented to O in accordance with the procedure common to the other psychophysical methods, offers obvious objections. How far the present results apply to Os who, though musical, are unable to recognize intervals readily, and to non-musical Os, is problematical. One may be reasonably safe in assuming, however, that for individuals who can recognize intervals these results are fairly representative. For such persons intermediate intervals, under attentive observation, would not usually be assimilated to their nearest 'norms' unless they were close to the liminal values of the latter.

[10] *Cf.* Stumpf & Meyer, *op. cit.*, 132, 147, 164 ff.

Summary

1. Judgments on musical intervals by the method of average error yield an average liminal value of 18 cents, an interval slightly less than a fifth of an equal semitone.

2. The 'psychophysical' interval has a range of variability, then, of about 40 cents. (Adjacent intervals of the tempered scale are 100 cents apart.)

3. Intervals which lie between those in ordinary use are not usually assimilated to their nearest 'norms' unless they are close to the liminal values of the latter. There is a range of about half an equal semitone midway between each tempered interval within which an interval would be recognized by O as being *neither* of the familiar intervals next above and below it.

4. The order of intervals, ranked according to precision of judgment, is approximately the same as that for degrees of fusion.

PSYCHOLOGICAL REVIEW PUBLICATIONS

Original contributions and discussions intended for the Psychological Review should be addressed to

Dr. John B. Watson, Editor PSYCHOLOGICAL REVIEW,
244 Madison Avenue, New York

Original contributions and discussions intended for the Journal of Experimental Psychology should be addressed to

Professor Madison Bentley,
Editor JOURNAL OF EXPERIMENTAL PSYCHOLOGY
University of Illinois, Urbana, Ill.

Contributions intended for the Psychological Monographs should be addressed to

Dr. S. I. Franz, Editor PSYCHOLOGICAL MONOGRAPHS,
University of California, Southern Branch, Los Angeles, California

Reviews of books and articles intended for the Psychological Bulletin, announcements and notes of current interest, and *books offered for review* should be sent to

Professor S. W. Fernberger, Editor PSYCHOLOGICAL BULLETIN,
University of Pennsylvania, Philadelphia

Titles and reprints intended for the Psychological Index should be sent to

Professor Walter S. Hunter, Editor PSYCHOLOGICAL INDEX,
Clark University, Worcester, Mass.

All business communications should be addressed to

Psychological Review Company, Princeton, New Jersey

Directory of American Psychological Periodicals

American Journal of Psychology—Ithaca, N. Y.; Cornell University.
Subscription $6.50. 624 pages ann. Edited by M. F. Washburn, K. M. Dallenbach, M. Bentley and E. G. Boring. Quarterly. General and experimental psychology. Founded 1887.

The Pedagogical Seminary and Journal of Genetic Psychology—Worcester, Mass.; Clark Univ. Sub. $7. 700 pages ann. Edited by Carl Murchison and an international coöperating board. Quarterly. Child behavior differential and genetic psychology. Founded 1891.

Psychological Review—Princeton, N. J.; Psychological Review Company.
Subscription $5.00. 480 pages annually.
Bi-monthly. General. Founded 1894. Edited by Howard C. Warren and John B. Watson.

Psychological Bulletin—Princeton, N. J.; Psychological Review Company.
Subscription $5.50. 720 pages annually. Psychological literature.
Monthly. Founded 1904. Edited by Samuel W. Fernberger.

Psychological Monographs—Princeton, N. J.; Psychological Review Company.
Subscription $6.00 per vol. 500 pp. Founded 1895. Edited by Shepherd I. Franz.
Published without fixed dates, each issue one or more researches.

Psychological Index—Princeton, N. J.; Psychological Review Company.
Subscription $2.00. 200 pp. Founded 1895. Edited by Walter S. Hunter.
An annual bibliography of psychological literature.

Journal of Philosophy—New York; 515 W. 116thSt.
Subscription $4. 728 pages per volume. Founded 1904.
Bi-weekly. Edited by F. J. E. Woodbridge, Wendell T. Bush and H. W. Schneider.

Archives of Psychology—Sub-Station 84, N. Y.; Archives of Psychology.
Subscription $5. 500 pp. per vol. Founded 1906. Edited by R. S. Woodworth.
Published without fixed dates, each number a single experimental study.

Journal of Abnormal and Social Psychology—Albany, N. Y.
Sub. $5. Boyd Printing Co. Edited by Morton Prince, in cooperation with Floyd H. Allport. Quarterly. 432 pages annually. Founded 1906. Abnormal and social.

Psychological Clinic—Philadelphia; Psychological Clinic Press.
Subscription $2.50. 288 pages. Edited by Lightner Witmer. Founded 1907.
Without fixed dates (9 numbers). Orthogenics, psychology, hygiene.

Training School Bulletin—Vineland, N. J.; The Training School.
Subscription $1. 160 pages ann. Edited by E. R. Johnstone. Founded 1904.
Monthly (10 numbers). Psychology and training of defectives.

Comparative Psychology Monographs—Baltimore; Williams and Wilkins Co.
Subscription $5. 500 pages per volume. Edited by W. S. Hunter.
Published without fixed dates, each number a single research.

Psychoanalytic Review—Washington, D. C.; 3617 10th St., N. W.
Subscription $6. 500 pages annually. Psychoanalysis.
Quarterly. Founded 1913. Edited by W. A. White and S. E. Jelliffe.

Journal of Experimental Psychology—Princeton, N. J.; Psychological Review Company.
Subscription $5.00. 480 pages annually. Experimental.
Bi-monthly. Founded 1916. Edited by Madison Bentley.

Journal of Applied Psychology—Bloomington, Ind.; Indiana University Press.
Subscription $4. 400 pages annually. Founded 1917.
Quarterly. Edited by James P. Porter and William F. Book.

Journal of Comparative Psychology—Baltimore; Williams and Wilkins Company.
Subscription $5. 500 pages annually. Founded 1921.
Bi-monthly. Edited by Knight Dunlap and Robert M. Yerkes.

Genetic Psychology Monographs—Worcester, Mass.; Clark University.
Subscription $7.00. 600 pp. ann. Edited by Carl Murchison.
Bi-monthly. Each number a complete research. Child behavior, differential and genetic psychology. Founded 1925.

Ex.P.